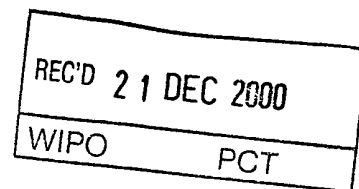


BUNDEREPUBLIK DEUTSCHLAND

01/02
(404)

EP 00108129

Prioritätsbescheinigung über die Einreichung
einer Patentanmeldung

4

Aktenzeichen: 199 39 578.0

Anmeldetag: 20. August 1999

Anmelder/Inhaber: Biosyn Arzneimittel GmbH, Fellbach/DE

Bezeichnung: Nukleinsäuremolekül, umfassend eine für ein
Hämocyanin kodierende Nukleinsäuresequenz

IPC: C 07 K, A 61 K, C 12 Q

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 30. November 2000
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Ebert

Nukleinsäuremolekül, umfassend eine für ein Hämocyanin kodierende Nukleinsäuresequenz

Die vorliegende Erfindung betrifft ein Nukleinsäuremolekül, umfassend eine für ein Hämocyanin, eine Hämocyanin-Domäne oder ein Fragment mit den immunologischen Eigenschaften wenigstens einer Domäne von Hämocyanin kodierende Nukleinsäuresequenz, diese umfassende Konstrukte, die Nukleinsäuresequenzen oder die Konstrukte umfassende Wirtszellen, Verfahren zum Herstellen von Hämocyanin-Polypeptiden und rekombinante Hämocyanin-Polypeptide.

Hämocyanin ist ein blaues Kupferprotein, das frei gelöst im Blut zahlreicher Mollusken und Arthropoden auftritt und den Sauerstoff transportiert. Von den Mollusken enthalten die Cephalopoden, Chitonen, die meisten Gastropoden sowie einige Bivalvia Hämocyanin. Hämocyanin ist bei den Arthropoden typisch für Arachniden, Xiphosuren, malakostrake Crustaceen und *Scutigera*. Zahlreiche Insektenarten weisen Proteine auf, die sich von Hämocyanin ableiten. Hämocyanine liegen extrazellulär vor und flottieren in der Hämolymphe.

Während das Arthropoden-Hämocyanin bei elektronenmikroskopischer Untersuchung einen Durchmesser von maximal 25 nm hat und eine Untereinheit ein Molekulargewicht von 75.000 Da aufweist, sind Molluskencyanine viel größer. So hat z.B. das Hämocyanin von *Megathura* einen Durchmesser von 35 nm und ist aus 2 Untereinheiten zusammengesetzt. Jede Untereinheit hat ein Molekulargewicht von ca. 400.000 Da und ist in acht sauerstoffbindende Domänen aufgeteilt, die jeweils ein Molekulargewicht von ca. 50.000 Da haben. Die Domänen unterscheiden sich immunologisch. Diese Domänen können durch limitierte Proteolyse aus der Untereinheit freigesetzt werden.

Das im Elektronenmikroskop sichtbare Hämocyanin der Gastropoden hat ein Molekulargewicht von ca. 8 Mio. Da und ist ein Di-Dekamer. Im Gegensatz hierzu ist das Hämocyanin der Cephalopoden als isoliertes Dekamer angeordnet, das sich auch in der Quartärstruktur deutlich vom Hämocyanin der Gastropoden unterscheidet.

Von besonderem immunologischen Interesse ist das Hämocyanin der kalifornischen Schlüssellochschnecke *Megathura crenulata*, einer „Keyhole Limpet“. Das Hämocyanin wird deshalb auch als Keyhole Limpet Hemocyanin (KLH) bezeichnet. Hämocyanine sind sehr starke Antigene. Die Immunisierung eines Vertebraten führt zu einer bisher wenig verstandenen, unspezifischen Aktivierung des Immunsystems. Durch die allgemeine Aktivierung des Immunsystems ist es dann möglich, auch eine Immunreaktion gegenüber anderen, bisher tolerierten Fremdstrukturen zu erreichen. KLH wird vor allem als Hapten-Träger verwendet, um so die Bildung von Antikörpern gegen das Hapten zu erreichen.

Neben *Megathura crenulata* gehört auch das Seeohr *Haliotis tuberculata* zur im Hinblick auf die Evolution relativ alten Gruppe der Archaeogastropoda. Es ist bekannt, daß auch *Haliotis* Hämocyanin produziert.

KLH ist ein Gemisch aus zwei unterschiedlichen Hämocyaninen, die als KLH1 und KLH2 bezeichnet werden. Die Untereinheit des KLH1 ist ein 390 kDa Polypeptid, das aus acht globulären Domänen besteht, die entsprechend ihrer Reihenfolge in der Untereinheit mit 1 a bis 1 h bezeichnet werden. KLH2 hingegen weist ein Molekulargewicht von 350 kDa auf und enthält nach neuesten Daten ebenfalls 8 Domänen, die als 2 a bis 2 h bezeichnet werden. *In vivo* bildet jede Art von Untereinheit Homo-Oligomere, wohingegen Hetero-Oligomere nicht beobachtet wurden.

Durch limitierte Proteolyse und gekreuzte Immunelektrophorese der Untereinheit von KLH1 und KLH2 wurden amino-, interne und carboxy-terminale Domänen erhalten, deren amino-terminale Sequenz bestimmt wurde (Söhngen et al., Eur. J. Biochem. 248 (1997), 602-614; Gebauer et al., Zoology 98(1994), 51-68). Die erhaltenen Sequenzen erlauben jedoch nicht den Entwurf sequenzspezifischer Primer und/oder Sonden, die für eine Hybridisierung mit genomischer DNA Erfolg versprechen. Obwohl beide KLH-Typen seit 1991 bzw. 1994 bekannt sind, konnte daher bisher keine Primärstruktur aufgeklärt werden.

Auf DNA-Ebene ist bisher in bezug auf Mollusken nur die cDNA-Sequenz der Häemocyanin-Untereinheit aus dem Cephalopoden *Octopus dofleini* bekannt (Miller et al., J. Mol. Biol. 278 (1998), 827-842). *Octopus dofleini* ist phylogenetisch von den Archæogastropoden sehr weit entfernt. Eine Häemocyanin-Gensequenz aus Mollusken ist bisher überhaupt nicht bekannt.

Wie von Miller et al. supra, beschrieben, ist es sowohl schwierig, eine einzige funktionelle Domäne (Funktionelle Einheit = Domäne; auch „funktionelle Domäne“ genannt) zu isolieren als auch Gewebe zu erhalten, das zur Aufreinigung von mRNA für die cDNA-Sequenzierung geeignet ist.

Bei der Analyse des Häemocyanins aus *Megathura crenulata* besteht eine weitere Schwierigkeit darin, daß die Versuchstiere ein Alter von 4 bis 8 Jahren erreicht haben müssen, um ihnen erstmals Hämolymphe entnehmen zu können. Nach Entnahme der Hämolymphe wird Häemocyanin bei diesen Tieren nicht nachproduziert. Bisher ist nicht bekannt, wie die Häemocyaninsynthese stimuliert werden könnte. Darüber hinaus ist die Zucht von *Megathura* äußerst aufwendig, da hierfür spezielle Strömungsbecken erforderlich sind.

Es ist daher eine Aufgabe der vorliegenden Erfindung, Mittel und Wege bereitzustellen, um Häemocyanin und/oder Domänen davon in ausreichender Menge und kostengünstig produzieren zu können. Dies umfaßt die weitere Aufgabe, ein Verfahren anzugeben, mit dem dieses Häemocyanin hergestellt werden kann.

Diese Aufgabe wird erfindungsgemäß gelöst durch ein Nukleinsäure-Molekül, umfassend eine für ein Häemocyanin, eine Häemocyanin-Domäne oder ein funktionelles Fragment davon mit den immunologischen Eigenschaften wenigstens einer Domäne eines Häemocyanins kodierende Nukleinsäuresequenz, wobei die Nukleinsäuresequenz ausgewählt ist aus

- (a) der Gruppe der nachfolgend angegebenen DNA-Sequenzen bzw. der ihnen entsprechenden RNA-Sequenzen:

in SEQ ID NO:1 (HtH1 Domäne a + Signalpeptid),

SEQ ID NO:2 (HtH1 Domäne b),

SEQ ID NO:3 (HtH1 Domäne c),

SEQ ID NO:4 (HtH1 Domäne d),

SEQ ID NO:5 (HtH1 Domäne e),

SEQ ID NO:6 (HtH1 Domäne f),

SEQ ID NO:7 (HtH1 Domäne g),

SEQ ID NO: 8 (HtH1 Domäne h),

SEQ ID NO:9 (partielle HtH2 Domäne b),

SEQ ID NO:10 (HtH2 Domäne c),

SEQ ID NO:11 (HtH2 Domäne d),

SEQ ID NO:12 (HtH2 Domäne e),

SEQ ID NO:13 (HtH2 Domäne f),

SEQ ID NO:14 (HtH2 Domäne g),

SEQ ID NO:15 (HtH2 Domäne h),

SEQ ID NO:16 (partielle KLH1 Domäne b),

SEQ ID NO:17 (KLH1 Domäne c),

SEQ ID NO:18 (KLH1 Domäne d),

~~SEQ ID NO:19 (partielle KLH1 Domäne e),~~

SEQ ID NO:20 (KLH2 Domäne b),

SEQ ID NO:21 (KLH2 Domäne c),

SEQ ID NO:22 (partielle KLH2 Domäne d),

SEQ ID NO:23 (KLH2 Domäne g),

SEQ ID NO:24 (partielle KLH2 Domäne h);

- (b) Nukleinsäuresequenzen, die mit dem Gegenstrang einer Nukleinsäuresequenz nach (a) hybridisieren und für ein Polypeptid kodieren, das die immunologischen Eigenschaften wenigstens einer Domäne eines Hämocyanins aufweist;
- (c) Nukleinsäuresequenzen, die aufgrund des genetischen Codes zu den unter (a) und (b) definierten DNA-Sequenzen degeneriert sind und für ein Polypeptid kodieren,

das die immunologischen Eigenschaften wenigstens einer Domäne eines Hä-mocyanins aufweist;

- (d) Nukleinsäuresequenzen, die mit einer der unter (a) bis (c) angegebenen Nuklein-säuresequenzen hybridisieren und deren Gegenstrang für ein Polypeptid kodiert, das die immunologischen Eigenschaften wenigstens einer Domäne eines Hä-mocyanins aufweist;
- (e) Nukleinsäuresequenzen, die wenigstens 60 % homolog zu einer der unter (a) ange-gebenen Nukleinsäuresequenzen sind;
- (f) Varianten der unter (a) bis (e) angegebenen Sequenzen, wobei die Varianten Addi-tionen, Deletionen, Insertionen oder Inversionen aufweisen und für ein Polypeptid kodieren, das die immunologischen Eigenschaften wenigstens einer Domäne von Hämocyanin aufweist; und
- (g) Kombinationen mehrerer der unter (a) bis (f) angegebenen DNA-Sequenzen.

Im nachfolgenden werden einige Begriffe näher erläutert, um klarzustellen, wie sie im Zusammenhang der vorliegenden Anmeldung verstanden werden sollen.

Der Begriff "Hämocyanin", so, wie er nachfolgend in der Beschreibung verwendet wird, umfaßt vollständiges Hämocyanin, Hämocyanin-Domänen und/oder Fragmente, Hä-mocyanin-Mutanten und Fusionsproteine. In bezug auf die Fusionsproteine sind insbe-sondere solche umfaßt, bei denen die Fusion Hämocyanin und Antigene umfaßt.

Unter „Domänen“ werden funktionelle Teilsequenzen der Hämocyanin-Untereinheiten verstanden, die beispielsweise durch limitierte Proteolyse voneinander abgetrennt wer-den können. Weiterhin können sie unterschiedliche immunologische Eigenschaften aufweisen.

Mit den „immunologischen Eigenschaften wenigstens einer Domäne von Hämocyanin“ ist die Eigenschaft eines Polypeptids gemeint, in gleicher Weise wie wenigstens eine Domäne von Hämocyanin eine immunologische Antwort des Empfängers zu induzieren, der mit dem Polypeptid immunisiert wird. Unter „immunologischer Antwort“ werden hier T- und/oder B-Zell-Antworten gegen Hämocyanin-Epitope verstanden, wie beispielsweise eine Antikörperproduktion. Die immunologische Reaktion kann beispielsweise beobachtet werden durch Immunisieren eines Säugers, wie z.B. einer Maus, einer Ratte oder eines Kaninchens mit dem entsprechenden Polypeptid und Vergleich der Immunantwort auf das zur Immunisierung verwendete Polypeptid mit der Immunantwort auf natürliche Hämocyanine.

Der Begriff "Antigen" umfaßt erfindungsgemäß sowohl Haptene, als auch schwache und starke Antigene. Haptene sind dadurch charakterisiert, daß sie Substanzen niedriger Molekülmasse (kleiner als 4000 Da) sind, jedoch ohne Kopplung an ein Trägermolekül nicht in der Lage sind, eine immunologische Reaktion auszulösen. Schwache Antigene sind Substanzen, die selbst bereits eine immunologische Reaktion auslösen können, deren Potential, eine immunologische Reaktion auslösen zu können, durch Kopplung mit einem Träger-Molekül auf Protein- und/oder DNA-Ebene, noch erhöht werden kann.

~~"His-Tag" bedeutet eine Sequenz von wenigstens 6 Histidin-Aminosäuren, die durch entsprechende Klonierung und Fusion mit einer exprimierbaren Sequenz zu einem Fusionsprotein mit wenigstens 6 His-Resten am NH₂-Terminus führt, das leicht durch Komplexierung mit einer Ni²⁺-Säule aufgereinigt werden kann.~~

"Klonierung" soll alle im Stand der Technik bekannten Klonierungsmethoden umfassen, die hier zum Einsatz kommen könnten, die jedoch nicht alle im einzelnen beschrieben werden, weil sie zum selbstverständlichen Handwerkszeug des Fachmanns gehören.

„Varianten“ einer Nukleinsäure umfassen Additionen, Deletionen, Insertionen oder Inversionen und kodieren für ein Polypeptid, das die immunologischen Eigenschaften wenigstens einer Domäne von Hämocyanin aufweist. Varianten können künstlich oder natürlich sein. Ein Beispiel für natürliche Varianten stellen allelische Varianten dar.

Unter "rekombinanter Expression in einer geeigneten Wirtszelle" sollen alle im Stand der Technik bekannten Expressionsmethoden in bekannten Expressionssystemen verstanden werden, die hier zum Einsatz kommen könnten, jedoch nicht alle im einzelnen beschrieben werden, weil sie zum selbstverständlichen Handwerkszeug des Fachmanns gehören.

Die im erfindungsgemäßen Nukleinsäuremolekül enthaltene Nukleinsäuresequenz kann genomische DNA, cDNA oder synthetische DNA sein, wobei unter synthetischen DNA-Sequenzen auch solche verstanden werden, die modifizierte Internukleosid-Bindungen enthalten. Weiter kann es sich bei den Nukleinsäuresequenzen um RNA-Sequenzen handeln, was z.B. für die Expression mittels rekombinanter Vektorsysteme erforderlich sein kann. Die Nukleinsäuresequenzen gemäß (b) sind beispielsweise erhältlich durch Verwenden einer nachweisbar markierten Sonde, die einer der unter (a) angegebenen Sequenzen oder einem Fragment bzw. deren Gegenstrang entspricht, zum Screening von cDNA-/genomischen DNA-Bibliotheken aus Mollusken oder Arthropoden. Die der cDNA-Bibliothek zugrundeliegende mRNA ist vorzugsweise aus Mollusken-Geweben zu erhalten, die Hämocyanin besonders stark exprimieren, wie z.B. Mantel-Gewebe aus Gastropoden und Branchialdrüsengewebe aus Cephalopoden.

Die Identifizierung positiver cDNA-/genomischer DNA-Klone erfolgt gemäß Standardverfahren. Vgl. Maniatis et al., Molecular Cloning (1989) Cold Spring Harbor Laboratory Press.

In einer bevorzugten Ausführungsform wird die unter (b) oder (d) angegebene Hybridisierung unter stringenten Bedingungen durchgeführt. Stringente Hybridisierungsbedingungen sind z.B. 68°C über Nacht in 0,5 x SSC; 1% Blockierungsreagenz (Boehringer Mannheim); 0,1 % Natriumlaurylsarcosinat und nachfolgendem Waschen mit 2 x SSC; 0,1 % SDS.

In einer bevorzugten Ausführungsform werden Nukleinsäuresequenzen bereitgestellt, die wenigstens 60 % homolog zu einer der unter (a) angegebenen Nukleinsäurese-

quenzen sind. Bevorzugt sind die Nukleinsäuresequenzen wenigstens 80 % homolog zu einer der unter (a) angegebenen Nukleinsäuresequenzen. Besonders bevorzugt sind die Nukleinsäuresequenzen wenigstens 90 % homolog zu einer der unter (a) angegebenen Nukleinsäuresequenzen. Insbesondere sind die Nukleinsäuresequenzen wenigstens 95 % homolog zu einer der unter (a) angegebenen Nukleinsäuresequenzen.

Erfindungsgemäß bedeutet der Ausdruck „Homologie“ Homologie auf DNA-Ebene, die gemäß bekannter Verfahren, z.B. der computergestützten Sequenzvergleiche (Basic local alignment search tool, S.F. Altschul et al., J. Mol. Biol. 215 (1990), 403-410) bestimmt werden kann.

Der dem Fachmann bekannte Ausdruck „Homologie“ bezeichnet den Grad der Verwandtschaft zwischen zwei oder mehr Nukleinsäuremolekülen, der durch die Übereinstimmung zwischen den Sequenzen bestimmt wird. Der Prozentsatz der „Homologie“ ergibt sich aus dem Prozentsatz identischer Bereiche in zwei oder mehr Sequenzen unter Berücksichtigung von Lücken oder anderen Sequenzbesonderheiten.

Die Homologie miteinander verwandter Nukleinsäuremoleküle kann mit Hilfe bekannter Verfahren bestimmt werden. In der Regel werden spezielle Computerprogramme mit den besonderen Anforderungen Rechnung tragenden Algorithmen eingesetzt.

Bevorzugte Verfahren zur Bestimmung der Homologie erzeugen zunächst die größte Übereinstimmung zwischen den untersuchten Sequenzen. Computerprogramme zur Bestimmung der Homologie zwischen zwei Sequenzen umfassen, sind jedoch nicht eingeschränkt auf, das GCG Programmpaket, einschließlich GAP (Devereux, J., et al., Nucleic Acids Research 12 (12): 387 (1984); Genetics Computer Group University of Wisconsin, Madison, (WI)); BLASTP, BLASTN und FASTA (Altschul, S. et al., J. Molec Biol 215:403/410 (1990)). Das BLAST X Programm kann vom National Centre for Biotechnology Information (NCBI) und aus weiteren Quellen bezogen werden (BLAST Handbuch, Altschul S., et al., NCB NLM NIH Bethesda MD 20894; Altschul, S., et al., J. Mol. 215:403/410 (1990)). Auch der bekannte Smith Waterman-Algorithmus kann zur Bestimmung von Homologien verwendet werden.

Bevorzugte Parameter für den Nukleinsäuresequenz-Vergleich umfassen die nachstehenden:

Algorithmus:	Needleman und Wunsch, J. Mol. Biol 48:443-453 (1970)
Vergleichsmatrix:	Übereinstimmung (matches) = + 10, Nichtübereinstimmung (mismatch) = 0
Lücken-Wert (Gap Penalty):	50
Lückenlängen-Wert: (Gap Length Penalty):	3

Das GAP-Programm ist auch zur Verwendung mit den vorstehenden Parametern geeignet. Die vorstehenden Parameter sind die Fehler-Parameter (default parameters) für Nukleinsäuresequenz-Vergleiche.

Weitere beispielhafte Algorithmen, Lücken-Öffnungs-Werte (gap opening penalties), Lückenausdehnungs-Werte (gap extension penalties), Vergleichsmatrizen einschließlich der im Programm-Handbuch, Wisconsin-Paket, Version 9, September 1997, genannten können verwendet werden. Die Auswahl wird von dem durchzuführenden Vergleich abhängen und weiterhin davon, ob der Vergleich zwischen Sequenzpaaren, wobei GAP oder Best Fit bevorzugt sind, oder zwischen einer Sequenz und einer umfangreichen Sequenz-Datenbank, wobei FASTA oder BLAST bevorzugt sind, durchgeführt wird.

Eine mit dem oben genannten Algorithmus ermittelte Übereinstimmung von 60 % wird im Rahmen dieser Anmeldung als 60 % Homologie bezeichnet. Entsprechendes gilt für höhere Homologiegrade.

In einer bevorzugten Ausführungsform ist die erfindungsgemäße DNA-Sequenz eine Kombination mehrerer der unter (a) bis (f) angegebenen DNA-Sequenzen, die durch dem Fachmann bekannte Fusion und gegebenenfalls Klonierung erhalten werden können. Diese Kombination sind von besonderem Interesse, da sie besonders immunogen sind. Insbesondere sind Kombinationen bevorzugt, die mehrere oder alle Domänen in der in der Untereinheit natürlicherweise vorkommenden Reihenfolge (a bis h) aufweisen.

Besonders bevorzugt sind dabei Ausführungsformen, in denen die für die Domänen kodierenden Nukleinsäuresequenzen direkt im Raster aneinandergeschlüsselt sind.

Weiterhin werden Konstrukte bereitgestellt, die die erfindungsgemäßen Nukleinsäuremoleküle umfassen. In einer bevorzugten Ausführungsform umfaßt das erfindungsgemäße Konstrukt einen zur Expression geeigneten Promotor, wobei die Nukleinsäuresequenz unter der Kontrolle des Promotors steht. Die Wahl des Promotors hängt vom zur Expression verwendeten Expressionssystem ab. Generell sind konstitutive Promotoren bevorzugt, jedoch sind auch induzierbare Promotoren wie z.B. der Metallothionein-Promotor möglich.

In einer weiterhin bevorzugten Ausführungsform umfaßt das Konstrukt ferner eine Antigen-kodierende Nukleinsäuresequenz, die direkt mit der erfindungsgemäßen Hämocyanin-Nukleinsäure verbunden ist. Die Antigen-kodierende Sequenz kann sowohl 5' als auch 3' relativ zur Hämocyanin-Sequenz oder auch an beiden Enden gelegen sein. Sie schließt im gleichen Leseraster entweder unmittelbar an die Hämocyanin-Sequenz an oder ist durch einen Nukleinsäure-Linker unter Wahrung des Leserasters mit ihr verbunden. Durch die Fusion der Antigen-kodierenden Sequenz mit der Hämocyanin-Sequenz ist die Bildung eines Fusionsproteins beabsichtigt, in dem die Antigen-kodierende Sequenz kovalent mit der Hämocyanin-Sequenz verbunden ist. Das erfindungsgemäße Antigen ist hierbei ein medizinisch relevantes Antigen, das beispielsweise ausgewählt ist aus: Tumorantigenen, Virusantigenen und Antigenen bakterieller oder parasitärer Pathogene. Tumorantigene können hierbei beispielsweise Rb und p53 sein. Vorzugsweise stammen die Virusantigene aus immunologisch relevanten Viren, wie z.B. Influenza-Virus, Hepatitis-Virus und HIV. Pathogenantigene sind unter anderem solche aus Säugerpathogenen, insbesondere humanpathogenen Organismen, wie z.B. Plasmodium. Bakterielle Antigene können z.B. von *Klebsiella*, *Pseudomonas*, *E. coli*, *Vibrio cholerae*, *Chlamydia*, *Streptococci* oder *Staphylococci* stammen.

In einer weiteren bevorzugten Ausführungsform umfaßt das Konstrukt ferner wenigstens ein Teil eines Vektors, insbesondere regulatorische Regionen, wobei der Vektor ausge-

wählt ist aus : Bacteriophagen wie λ -Derivaten, Adenoviren, Vacciniaviren, Baculoviren, SV40-Viren und Retroviren, vorzugsweise MoMuLV (Moloney Murine Leukemia Virus).

Ferner ist ein Konstrukt bevorzugt, das zusätzlich eine His-Tag-kodierende DNA-Sequenz umfaßt, die bei Expression des Konstrukts zur Bildung eines Fusionsproteins mit einem His-Tag am NH_2 -Terminus des Hämocyanins führt, welches die Aufreinigung des Proteins an einer Nickel-Säule durch Chelat-Bildung erleichtert.

Ferner stellt die Erfindung Wirtszellen bereit, die das Konstrukt enthalten und die zur Expression des Konstruktes geeignet sind. Im Stand der Technik sind zahlreiche prokaryontische und eukaryontische Expressionssysteme bekannt, wobei die Wirtszellen beispielsweise ausgewählt sind aus prokaryontischen Zellen wie *E. coli* oder *B. subtilis*, aus eukaryontischen Zellen wie Hefezellen, Insektenzellen und Säugerzellen, z.B. CHO-Zellen, COS-Zellen oder HeLa-Zellen, sowie Derivaten davon. Im Stand der Technik sind beispielsweise bestimmte CHO-Produktionslinien bekannt, deren Glykosylierungsmuster im Vergleich zu CHO-Zellen verändert sind. Die durch die Verwendung Glykosylierungs-defizienter oder Glykosylierungs-verringelter Wirtszellen erhaltenen Hämocyanine verfügen möglicherweise über zusätzliche Epitope, die bei vollständiger Glykosylierung ansonsten dem Immunsystem des Empfängers nicht zugänglich sind, so daß Hämocyanine mit verringerter Glykosylierung unter Umständen eine erhöhte Immunogenität aufweisen.

Gegenstand der vorliegenden Erfindung ist weiter ein Verfahren zum Herstellen eines Hämocyanin-Polypeptids. Dazu wird das erfindungsgemäße Nukleinsäuremolekül und/oder das Konstrukt in einer geeigneten Wirtszelle exprimiert und das Protein aus der Wirtszelle oder dem Medium mittels üblicher Verfahren isoliert.

Dem Fachmann sind zahlreiche Verfahren zur Expression von DNA-Sequenzen bekannt; vergleiche Recombinant Gene Expression Protocols in Methods in Molecular Biology, Band 62, Humana Press Totowa New Jersey (1995). Die Expression kann sowohl konstitutiv als auch induzierbar sein, wobei Induktoren wie beispielsweise IPTG und Zn^{2+} dem Fachmann bekannt sind. Das hergestellte Hämocyanins kann, falls ein His-

Tag an den NH_2 -Terminus des Hämocyanin fusioniert wurde, durch Chelat-Bildung an einer Nickel-Säule aufgereinigt werden. Verfahren zum Aufreinigen von Hämocyanin, insbesondere KLH, finden sich in Harris et al., Micron 26 (1995), 201-212. Vorzugsweise wird das Hämocyanin durch Ionenaustausch-Chromatographie und/oder Gelfiltrationschromatographie aufgereinigt. Die Durchführung dieser Maßnahmen ist dem Fachmann bekannt.

In einer weiteren bevorzugten Ausführungsform wird das erfindungsgemäß hergestellte Hämocyanin modifiziert. Die Modifikationen umfassen hierbei die Di-, Oligo- und Polymerisierung des monomeren Ausgangsprodukts beispielsweise durch Quervernetzung, z.B. durch Dicyclohexylcarbodiimid oder Pegylierung oder Assoziation (self assembly). Die somit hergestellten Di-, Oligo- und Polymere können voneinander durch Gelfiltration abgetrennt werden. Insbesondere beabsichtigt ist die Bildung von Dekameren, Didekameren oder Multi-Dekameren. Weitere Modifikationen umfassen Seitenketten-Modifikationen, beispielsweise von ϵ -Amino-Lysin-Resten des Hämocyanins, oder Amino- bzw. Carboxy-terminale Modifikationen. Besonders bevorzugt ist die Modifikation des Hämocyanins durch kovalente Bindung an ein Antigen, wobei das Antigen stöchiometrisch oder nicht-stöchiometrisch mit dem Hämocyanin umgesetzt sein kann. Das Antigen ist vorzugsweise ausgewählt aus Tumorantigenen, Virusantigenen und Pathogenantigenen wie oben ausgeführt. Weitere Modifikationen umfassen posttranslationale Ereignisse, z.B. die Glykosylierung oder die partielle oder vollständige Deglykosylierung des Proteins.

In einer bevorzugten Ausführungsform ist das erhaltene Hämocyanin bei rekombinanter Expression in Prokaryonten oder Glykosylierungs-defizienten Eukaryonten nicht-glykosyliert. Ebenfalls in Betracht gezogen wird erfindungsgemäß Hämocyanin, das durch rekombinante Expression in zur Glykosylierung fähigen Eukaryonten wie Hefezellen, Insektenzellen oder Säugerzellen, wie CHO-Zellen oder HeLa-Zellen, glykosyliert ist.

In einer weiteren Ausführungsform werden Hämocyanin-Polypeptide zur Verfügung gestellt, die eine Aminosäuresequenz umfassen, wobei die Aminosäuresequenz von einer oder mehreren der erfindungsgemäßen Nukleinsäuremoleküle kodiert wird.

Bevorzugt werden Hämocyanin-Polypeptide zur Verfügung gestellt, die wenigstens eine aus der folgenden Gruppe ausgewählte Aminosäuresequenz umfassen:

SEQ ID NO:25 (HtH1 Domäne a + Signalpeptid),

SEQ ID NO:26 (HtH1 Domäne b),

SEQ ID NO:27 (HtH1 Domäne c),

SEQ ID NO:28 (HtH1 Domäne d),

SEQ ID NO:29 (HtH1 Domäne e),

SEQ ID NO:30 (HtH1 Domäne f),

SEQ ID NO:31 (HtH1 Domäne g),

SEQ ID NO:32 (HtH1 Domäne h),

SEQ ID NO:33 (partielle HtH2 Domäne b),

SEQ ID NO:34 (HtH2 Domäne c),

SEQ ID NO:35 (HtH2 Domäne d),

SEQ ID NO:36 (HtH2 Domäne e),

SEQ ID NO:37 (HtH2 Domäne f),

SEQ ID NO:38 (HtH2 Domäne g),

SEQ ID NO:39 (HtH2 Domäne h),

SEQ ID NO:40 (partielle KLH1 Domäne b),

SEQ ID NO:41 (KLH1 Domäne c),

SEQ ID NO:42 (partielle KLH1 Domäne d),

SEQ ID NO:43 (partielle KLH1 Domäne e),

SEQ ID NO:44 (KLH2 Domäne b),

SEQ ID NO:45 (KLH2 Domäne c),

SEQ ID NO:46 (partielle KLH2 Domäne d),

SEQ ID NO:47 (KLH2 Domäne g),

SEQ ID NO:48 (partielle KLH2 Domäne h),

oder ein Fragment einer dieser Sequenzen, das die immunologischen Eigenschaften wenigstens einer Domäne von Hämocyanin aufweist.

In einer weiteren Ausführungsform stellt die Erfindung Hämocyanin-Polypeptide, erhältlich durch das rekombinante Herstellungsverfahren oder Modifikationen davon, bereit.

Bevorzugt sind Hämocyanin-Polypeptide, die jede der Sequenzen SEQ ID NO: 25 bis 32 umfassen, und Hämocyanin-Polypeptide, die die Sequenzen SEQ ID NO: 33 bis 39 umfassen. Besonders bevorzugt handelt es sich bei diesen Hämocyanin-Polypeptiden um Hämocyanin 1 oder 2 aus *Haliotis tuberculata*.

Insbesondere bevorzugt ist Hämocyanin 1 aus *Haliotis tuberculata*, das ein scheinbares Molekulargewicht von 370 kDa in SDS-PAGE unter reduzierenden Bedingungen aufweist. Weiterhin ist insbesondere Hämocyanin 2 aus *Haliotis tuberculata* bevorzugt, das ein scheinbares Molekulargewicht von 370 kDa in SDS-PAGE PAGE unter reduzierenden Bedingungen aufweist. Die Hämocyanine sind durch das in den Beispielen beschriebene selektive Dissoziationsverfahren aus Gesamt-Hämocyanin aus *Haliotis tuberculata* erhältlich.

Weiterhin bevorzugt sind Hämocyanin-Polypeptide, die jede der Sequenzen SEQ ID NO: 40 bis 43 umfassen, und Hämocyanin-Polypeptide, die jede der Sequenzen SEQ ID NO: 44 bis 48 umfassen.

Besonders bevorzugt handelt es sich bei diesen Hämocyanin-Polypeptiden um vollständiges Hämocyanin 1 (KLH1) oder 2 (KLH2) aus *Megathura crenulata*.

Weiterhin wird nicht-glykosyliertes und glykosyliertes Hämocyanin-Polypeptid, erhältlich durch Expression in zur Glykosylierung fähigen bzw. unfähigen Wirtszellen, bereitgestellt. Je nach vorgesehener Verwendung des Hämocyanin-Polypeptids kann das Glykosylierungsmuster von Hefe, insbesondere methylotropher Hefe, von COS- oder HeLa-Zellen bevorzugt sein.

Die Erfindung betrifft weiterhin pharmazeutische Zusammensetzungen, die die erfindungsgemäßen Nukleinsäuremoleküle und physiologisch verträgliche Zusatzstoffe, die im Stand der Technik bekannt sind, enthalten. Vorzugsweise werden die pharmazeutischen Zusammensetzungen zur unspezifischen Immunstimulierung in Form einer Genterapie eingesetzt, wobei nach Transformation mit einem geeigneten Vektor Hämocyanin-Polypeptide exprimiert werden und zur Antigenisierung des Gewebes dienen.

Insbesondere sieht die vorliegende Erfindung die Verwendung eines erfindungsgemäßen Nukleinsäuremoleküls, das mit einer Antigen-kodierenden DNA-Sequenz verbunden ist, zur spezifischen Immunisierung gegen dieses Antigen vor. Die Immunisierung beruht hierbei, ohne an diese Theorie gebunden zu sein, auf der unspezifischen Stimulierung des Immunsystems durch Hämocyanin-Polypeptid-Epitope und die weitergehende spezifische Immunisierung durch Erkennung von Antigen-Epitopen durch das Immunsystem.

Eine solche Immunisierung ist besonders wertvoll im Hinblick auf Pathogen-Antigene, ganz besonders aber im Hinblick auf Tumorantigene. Die Anwendbarkeit der erfindungsgemäßen pharmazeutischen Zusammensetzung zur Behandlung von Tumorerkrankungen ergibt sich auch aus der Kreuzreaktivität der gebildeten Hämocyanin-spezifischen Antikörper mit Kohlenhydratresten, die auf der Oberfläche von Tumoren auftreten, wie z.B. dem Thomsen-Friedenreich-Antigen, das bei der Mehrzahl von humanen Tumoren wie Epithelialkarzinomen, Ovarialkarzinom, Kolonrektalkarzinom, Mammakarzinom, Bronchialkarzinom und Harnblasenkarzinom auftritt.

Weiterhin können die erfindungsgemäßen pharmazeutischen Zusammensetzungen zum Behandeln von parasitären Erkrankungen wie Schistosomiasis und für die Kokain-Mißbrauchsvorsorge eingesetzt werden.

Als weitere Ausführungsform der vorliegenden Erfindung werden pharmazeutische Zusammensetzungen zur Verfügung gestellt, die ein erfindungsgemäßes Hämocyanin-Polypeptid in Verbindung mit einem oder mehreren physiologisch verträglichen Zusatzmitteln enthalten. Wie oben bereits erwähnt, kann ein solches Hämocyanin-Poly-

peptid aus einer vollständigen Hämocyanin-Untereinheit, aus einer oder mehreren Domänen sowie aus einem oder mehreren Fragmenten solcher Domänen bestehen, vorausgesetzt, daß diese Fragmente noch die immunologischen Eigenschaften wenigstens einer Domäne eines Hämocyanins aufweist. Eine solche pharmazeutische Zusammensetzung eignet sich durch die entweder unspezifische Immunstimulation, die allein auf das Hämocyanin zurückzuführen ist, oder durch die spezifische Immunreaktion auf mit dem Hämocyanin assoziierte Antigene z.B. als Antiparasitenmittel, Antivirumittel oder Antitumormittel. So kann sie z.B. zum Behandeln von Schistosomiasis, Epithelialkarzinomen, Ovarialkarzinom, Kolonrektalkarzinom, Mammakarzinom, Bronchialkarzinom und Harnblasenkarzinomen eingesetzt werden, eignet sich jedoch auch zum Behandeln von Bluthochdruck. Die Behandlung von Bluthochdruck wird erreicht, indem eine Immunisierung mit Hilfe von erfindungsgemäßen Hämocyanin- β -adrenergen-Rezeptorpeptid-Konstrukten und/oder Fusionsproteinen durchgeführt wird.

In einer weiteren Ausführungsform werden die erfindungsgemäßen pharmazeutischen Zusammensetzungen als Impfstoffe verwendet. Sie können somit einen wertvollen Beitrag zur Prophylaxe von durch bekannte Pathogene verursachte Erkrankungen leisten. Dies gilt insbesondere für pharmazeutische Zusammensetzungen, in denen ein Hämocyanin-Polypeptid an ein Virus, Virusbestandteil, abgetötete Bakterien, Bakterienbestandteile, insbesondere Oberflächenproteine aus Virus- oder Bakterienhüllen, DNA, DNA-Bestandteile, anorganische oder organische Moleküle, z.B. Kohlenhydrate, Peptide und/oder Glykoproteine gebunden ist.

Gemäß einer weiteren bevorzugten Ausführungsform wird die erfindungsgemäße pharmazeutische Zusammensetzung zur Kokain-Mißbrauchsvorsorge verwendet.

Zur Applikation sowohl der erfindungsgemäßen Nukleinsäuremoleküle als auch der Hämocyanin-Polypeptide eignen sich insbesondere Liposomen. Dementsprechend umfaßt die vorliegende Erfindung Liposomen, die ein erfindungsgemäßes Nukleinsäuremolekül, ein erfindungsgemäßes Konstrukt oder ein erfindungsgemäßes Hämocyanin-Polypeptid umfassen.

Dem Fachmann sind verschiedene Verfahren zum Herstellen von Liposomen, die für pharmazeutische Zwecke verwendbar sind, bekannt. Die Selektivität der erfindungsgemäßen Nukleinsäuremoleküle oder Hämocyanin-Polypeptide enthaltenden Liposomen kann durch den zusätzlichen Einbau von Zellerkennungsmolekülen in die Liposomen erhöht werden, die selektiv an Zielzellen binden. Hierzu eignen sich insbesondere Rezeptorliganden, die an Rezeptoren der Zielzellen binden, oder, besonders im Fall von Tumoren, Antikörper, die gegen Oberflächenantigene der jeweils anvisierten Zielzellen gerichtet sind.

Die erfindungsgemäßen Hämocyanin-Polypeptide sind außerdem als Trägermolekül für Arzneistoffe, wie z. B. Cytostatika, vorgesehen. Die Vergrößerung des Molekulargewichtes verlängert die physiologische Halbwertszeit der Arzneistoffe erheblich, da der Verlust durch Ultrafiltration in der Niere deutlich verringert ist.

Die Zubereitung der Impfstoffe erfolgt nach dem Fachmann bekannten Verfahren; in einigen Ausführungsformen ist die zusätzliche Verwendung von Adjuvanzen wie z. B. Freundsches Adjuvanz oder Polysacchariden vorgesehen.

Die Erfindung stellt ferner Antikörper bereit, die spezifisch mit dem erfindungsgemäßen Hämocyanin-Polypeptid reagieren und erhältlich sind durch Immunisieren eines Versuchstieres mit einem Hämocyanin-Polypeptid. Polyklonale Antikörper können durch Immunisieren beispielsweise von Kaninchen und anschließendem Gewinnen von Antisera erhalten werden. Monoklonale Antikörper können gemäß Standardverfahren durch Immunisieren von z.B. Mäusen, Gewinnen und Immortalisieren der Milzzellen und Klonieren der Hybridome, die für Hämocyanin spezifische Antikörper produzieren, erhalten werden.

Weiterhin wird ein Screening-Verfahren zum Identifizieren von Tumor-spezifischer DNA in einer Zelle bereitgestellt, das die Schritte umfaßt:

a) das Inkontaktbringen zellulärer DNA und/oder zellulären Proteins mit einer Sonde umfassend das erfindungsgemäße Nukleinsäuremolekül und/oder den erfindungsgemäßen Antikörper und

b) das Nachweisen der spezifischen Bindung.

Vorzugsweise ist der nachzuweisende Tumor ein Harnblasenkarzinom, Epithelialkarzinom, Ovarialkarzinom, Mammakarzinom, Bronchialkarzinom oder Kolonrektalkarzinom.

Es ist beabsichtigt, mit den nachfolgenden Figuren und Beispielen die Erfindung zu erläutern, diese jedoch in keiner Weise einzuschränken. Dem Fachmann sind aufgrund der Beschreibung und der Beispiele weitere Ausführungsformen zugänglich, die ebenfalls umfaßt sind.

Fig. 1 zeigt die Charakterisierung und Aufreinigung von *Haliotis tuberculata* Hämocyanin (HtH):

- (a) Elektronenmikroskopie von negativ gefärbtem Gesamt-HtH, das durch Ultrazentrifugation von Zell-freier Hämolymphe aufgereinigt wurde;
- (b) SDS-Polyacrylamid-Gelelektrophorese (7,5 % Polyacrylamid) von HtH1 im Vergleich zu KLH (MW 370 kDa);

- (c) Native Polyacrylamid-Gelelektrophorese (5 % Polyacrylamid) der HtH-Untereinheiten-Präparation, wobei die Anode am unteren Rand liegt;
- (d) Gekreuzte Immunelektrophorese der beiden HtH-Untereinheiten unter Verwendung von anti-HtH-Antikörpern aus Kaninchen;
- (e) Elektronenmikroskopie der verbleibenden HtH1-Didekamere (weiße Pfeile) nach der selektiven Dissoziation von HtH2 (schwarze Pfeile);
- (f) Elutionsprofil der Gelfiltrationschromatographie (Biogel A15m) in Gegenwart von Ammoniummolybdat/Polyethylenglykol-Lösung (pH 5,9) nach der selektiven Dissoziation von HtH2 in seine Untereinheit und nachfolgender Anreicherung von HtH1 durch Ultrazentrifugation;

- (g) Native Polyacrylamid-Gelelektrophorese (6,5 % Polyacrylamid) von durch Gelchromatographie aufgereinigten HtH1- und HtH2-Untereinheiten im Vergleich zum Ausgangsmaterial;
- (h,i) Gekreuzte Immunelektrophorese von chromatographisch aufgereinigten HtH-Untereinheiten; und
- (j,m) Gekreuzte Immunelektrophorese der aufgereinigten HtH-Untereinheiten unter Verwendung von anti-KLH-Antikörpern aus Kaninchen, die für KLH1 bzw. KLH2 spezifisch sind.

Fig. 2 zeigt die Untersuchung der Untereinheiten-Organisation von HtH1, wobei für die Immunelektrophorese anti-HtH1-Antikörper aus Kaninchen verwendet wurden und die Anode sich auf der linken Seite befand:

- (a) Gekreuzte Immunelektrophorese nach limitierter Proteolyse von HtH1 mit Hilfe von Elastase;
- (b) SDS-Polyacrylamid-Gelelektrophorese (7,5 % Polyacrylamid) von Elastase-gespaltener HtH1-Untereinheit;
- (c,d,g-j,l,n,p) Gekreuzte Immunelektrophorese der Elastase-Spaltprodukte der HtH1-Untereinheit;
- (e) Gekreuzte Immunelektrophorese nach limitierter Proteolyse von HtH1 mit Hilfe von V8-Protease;
- (f) SDS-Polyacrylamid-Gelelektrophorese (7,5 % Polyacrylamid) von V8 Protease-gespaltener HtH1-Untereinheit und
- (k,m,o) Gekreuzte Immunelektrophorese nach limitierter Proteolyse von HtH1 mit Hilfe der drei angegebenen Proteasen.

Fig. 3 zeigt die Auftrennung von proteolytischen Spaltprodukten der Untereinheit HtH1 mit Hilfe von HPLC.

Fig. 4 zeigt die cDNA-Sequenz von HtH1 in Verbindung mit der Intronstruktur.

Fig. 5 zeigt die abgeleitete Primärstruktur von HtH1.

Fig. 6 zeigt die cDNA-Sequenz von Hth2 in Verbindung mit der Intronstruktur.

Fig. 7 zeigt die abgeleitete Primärstruktur von Hth2.

Fig. 8 zeigt die cDNA-Sequenz von KLH1 in Verbindung mit der Intronstruktur.

Fig. 9 zeigt die abgeleitete Primärstruktur von KLH1.

Fig. 10 zeigt die cDNA-Sequenz von KLH2 in Verbindung mit der Intronstruktur.

Fig. 11 zeigt die abgeleitete Primärstruktur von KLH2.

BEISPIELE

Material und Methoden

1. Gewinnen der Hämolymphe und Isolieren von Hämocyanin

Individuen des europäischen Seeohrs *Haliothis tuberculata* aus dem Bereich der französischen Atlantikküste wurden von S.M.E.L (Blainville sur Mer, Frankreich) und der Firma Biosyn (Fellbach, Deutschland) bereitgestellt. Die Tiere wurden in einem 300 l Seewasser-Aquarium bei 17 ° C gehalten und mit Braunalgen gefüttert. Zur Entnahme der Hämolymphe wurden die Seeohren in einem verschlossenen Plastiksack auf Eis gestellt. Nach einer Stunde waren große Volumina an Hämolymphe durch ihre Haut sezerniert worden. Es stellte sich heraus, daß das durch dieses Verfahren erhaltene Hämocyanin identisch ist mit dem Hämocyanin, das durch Einschneiden einer Mulde in den Fuß heruntergekühlter Meeresschnecken unter Verwendung einer Skalpellklinge gesammelt werden konnte. Die Blutzellen wurden von der Hämolymphe durch Zentrifugation bei 800 g über 30 min bei 4 ° C abgetrennt. Das gesamte Hämocyanin wurde sofort danach durch präparative Ultrazentrifugation bei 30000 g über 4 Stunden bei 4 ° C sedimentiert. Der Überstand wurde verworfen und das blaue Hämocyanin-Pellet wurde

über Nacht in "Stabilisierungs-Puffer" (0,05 M Tris, 5 mM CaCl_2 , 5 mM MgCl_2 , 0,15 M NaCl, 1 mM PMSF, pH 7,4) suspendiert und bei 4 ° C gelagert.

Intaktes HtH1 wurde unter Verwendung des von Harris et al., 1995, supra, beschriebenen Verfahrens aus dem gesamten HtH durch selektive Dissoziation von HtH2 in Ammoniummolybdat/Polyethylenglykol (1%/0,2%)-Lösung, pH 5,9 und nachfolgender Ultrazentrifugation erhalten. Das entstandene, teilweise aufgereinigte HtH1-Pellet wurde aufgelöst und durch Gelfiltration auf einer Biogel A15m-Vorrichtung zur Homogenität aufgereinigt. Der letzte Schritt ergab geringe Mengen von aufgereinigtem HtH2. Natives HtH1 und HtH2 wurde durch Dialyse gegen "Dissoziationspuffer" (0,13 M Glycin/NaOH, pH 9,6) bei 4 ° C über Nacht quantitativ in die Untereinheiten dissoziiert; die Gegenwart von EDTA war nicht erforderlich. 1 mM PMSF wurde bei jeder Stufe der Aufreinigung hinzugefügt, um die Proteolyse zu hemmen.

2. Elektronenmikroskopie

Konventionelles "negative staining" wurde mit dem Einzel-Tropfen-Verfahren (Harris und Horne in Harris, J.R. (Herausgeber) Electron microscopy in biology, (1991), IRL Press Oxford, S. 203-228) durchgeführt. Kohlenstoffträger-Filme wurde anfänglich über 20 Sekunden glühentladen, um sie hydrophil und für das Protein adsorptiv zu machen. Man lässt die Proteinproben an den Kohlenstofffilmen über 60 Sekunden adsorbieren. Sodann werden die Puffersalze durch sequentielles Waschen mit vier aufeinanderfolgenden 20 μl Wassertropfen entfernt. Die Gitter werden schließlich mit einem 20 μl Tropfen 5 % wässrigen Ammoniummolybdats, enthaltend 1 % Trehalose (pH 7,0) negativ gefärbt und bei Raumtemperatur trocknen gelassen. Für die elektronenmikroskopische Untersuchung wird ein Zeiss EM 900 Transmissionselektronenmikroskop verwendet.

3. Polyacrylamid-Gel-Elektrophorese und Immunelektrophorese

SDS-Polyacrylamid-Gel-Elektrophorese (SDS-PAGE) wurde gemäß des Verfahrens von Laemmli (Nature 227 (1970), 670-685) durchgeführt. Zur nativen PAGE wurde ein alkalisches System gemäß Markl et al. (1979) J. Comp. Physiol. 133 B, 167-175 mit einem

0,33 M Tris/Borat, pH 9,6 als Gelpuffer und 0,065 M Tris/Borat, pH 9,6 als Elektrodenpuffer verwendet. Gekreuzte und "crossed-line" Immunelektrophorese (IE) wurden gemäß Weeke (Scand. J. Immunol. 2 (1973), Suppl. 1, 47-56) bzw. Kroll (Scand. J. Immunol. 2, Suppl. 1 (1973), 79-81) durchgeführt. Kaninchenantikörper wurden von Charles River Deutschland (Kisslegg, Deutschland) gegen dissoziiertes Gesamt-HtH und aufgereinigtes HtH1 erzeugt. Das Immunisierungsverfahren wurde durchgeführt gemäß Markl und Winter (J. Comp. Physiol. 159B (1989), 139-151).

4. Limitierte Proteolyse und Isolierung der Fragmente

Die limitierte Proteolyse wurde bei 37 ° C in 0,13 M Glycin/NaOH, pH 9,6 durch Hinzufügen eines der nachstehenden Enzyme (Sigma, Deisenhofen, Deutschland), die in 0,1 M NH_4HCO_3 , pH 8,0 gelöst waren, durchgeführt: *Staphylococcus aureus* V8-Protease Typ XVII (8400), Papain Typ II aus Papaya-Milch (P-3125), Rinder-Pankreas-Elastase Typ IV (E-0258), Chymotrypsin und Trypsin. Die Hämocyanin-Konzentration lag zwischen 1 und 10 mg/ml. Die Endkonzentration des Enzyms betrug 2 % (Gewicht/Gewicht). Die Proteolyse wurde nach 5 Stunden durch Einfrieren auf -20 ° C beendet. Das HPLC-Verfahren wurde auf einer Vorrichtung von Applied Biosystems (BAI, Bensheim, Deutschland) durchgeführt, die mit einem Modell-1000S-Dioden-Array-Detektor ausgestattet war. Die proteolytischen Fragmente wurden auf kleine Mono-Q-Anionenaustauscher-Säule aufgetragen (Pharmacia, Freiburg, Deutschland), die mit 0,02 M Tris/HCl, pH 8,0 äquilibriert worden war und mit einem linearen Natriumchlorid-Gradienten (0,0 M - 0,5 M CaCl) im gleichen Puffer bei einer Flußrate von 1 ml/min eluiert wurden. Alternativ wurden die proteolytischen Fragmente durch Ausschneiden der Banden aus nativen PAGE-Gelen (Markl et al., 1979) J. Comp. Physiol. 133 B, 167-175 isoliert, nachdem sie zuvor mit dem Roti-White-System (Roth, Karlsruhe, Deutschland) gemäß Fernandez-Patron et al. (1995) Anal. Biochem. 224, 203-211 invers-gefärbt worden waren. Zur nachfolgenden Spaltung mit einem zweiten Enzym wurden die isolierten Fragmente zuerst Übernacht gegen 0,13 M Glycin/NaOH, pH 9,6 dialysiert, um NaCl zu entfernen.

5. Aminosäuresequenzanalyse

Die durch das HPLC-Verfahren erhaltenen Proteine wurden in SDS-enthaltendem Probenpuffer denaturiert und durch SDS-PAGE (Laemmli, 1970, supra; 7,5 % Polyacrylamid) aufgetrennt. Um die Blockierung des NH₂-Terminus zu verhindern wurde 0,6 % (Gewicht/Gewicht) Thioglykolsäure zum Kathodenpuffer (Walsh et al., Biochemistry 27 (1988), 6867-6876) hinzugefügt. Die Proteinbanden wurden durch Elektro-Transfer auf ProBlot-Membranen (Applied Biosystems, Deutschland) in einer vertikalen Blotting-Kammer übertragen (25 mM Borat-Puffer, pH 8,8, enthaltend 2 mM EDTA; 10 min/100 mA, 15 min/200 mA, 12 h/300 mA). Der Nachweis der einzelnen Polypeptide auf den Membranen wurde mit der Ponceau-S-Färbung durchgeführt. Die interessierenden Polypeptid-Banden wurden ausgeschnitten und in einer 477A-Protein-Sequenzier-Vorrichtung von Applied Biosystems sequenziert. Die Mengen der auf die Sequenzier-Vorrichtung aufgetragenen Polypeptide lag im unteren pmol-Bereich.

6. cDNA-Klonierung und Sequenzanalyse

Eine Lambda-cDNA-Expressionsbibliothek wurde aus Poly(A⁺)-RNA aus *Haliotis* Mantelgewebe unter Verwendung des Vektors Lambda ZAP Express[®] gemäß den Herstelleranweisungen (Stratagene, Heidelberg, Deutschland) hergestellt. Die Klone wurden unter Verwendung von HtH-spezifischen Kaninchenantikörpern isoliert. Die Nukleotidsequenzierung wurde auf beiden Strängen unter Verwendung des Taq Dye deoxy Terminator[®]-Systems durchgeführt. Die Anordnung der Sequenzen wurde mit der Software CLUSTAL W (1.7)[®] und TREEVIEW[®] (Thompson et al., Nucl. Acids Res. 22 (1994), 4673-4680) durchgeführt.

Beispiel 1:

Isolierung von HtH und Auftrennung zweier unterschiedlicher Typen (HtH1 und HtH2)

Die Hämolymphe wurde aus adulten Seeohren gewonnen. Die Blutzellen wurden durch Zentrifugation entfernt und das Hämocyanin wurde sodann durch Ultrazentrifugation sedimentiert. Das blaue Hämocyanin-Pellet wurde in "Stabilisierungspuffer" (pH 7,4) wieder aufgelöst und in der Elektronenmikroskopie untersucht (Figur 1a). Es bestand

hauptsächlich aus typischen Di-Dekameren, begleitet von einem kleinen Anteil an Dekameren und Tridekameren. Die Denaturierung in 2 % SDS in Gegenwart reduzierender Substanzen und anschließender SDS-PAGE-Trennung ergab eine einzige Bande, die dem Polypeptid mit einem scheinbaren Molekulargewicht von 370 kDa entsprach, welches nur geringfügig unterhalb des scheinbaren Untereinheiten-Gewichts von KLH (Figur 1b) liegt. Die vollständige Dissoziation der Oligomere wie der Di-Dekamere in die nativen Polypeptide (Untereinheiten) wurde erreicht durch Übernacht-Dialyse von HtH gegen "Dissoziationspuffer" (pH 9,6). Natives PAGE-Verfahren, das auf diese Proben angewendet wurde, ergab eine Haupt- und eine Neben-Komponente (Figur 1c). Die gekreuzte Immunelektrophorese (gekreuzte IE) unter Verwendung von gegen aufgereinigtes Gesamt-HtH-erzeugten polyklonalen Kaninchen-Antikörpern zeigte zwei Komponenten, die immunologisch unterschiedlich sind, jedoch die klassische Reaktion einer teilweise immunologischen Identität aufweisen (Figur 1d). Ihre präparative Isolierung (Figur 1e-i) zeigte, daß sie die Untereinheiten von zwei unterschiedlichen HtH-Typen darstellen, bezeichnet als HtH1 und HtH2, und die Muster der nativen PAGE- und gekreuzten IE-Verfahren konnten jeder einzelnen zugeordnet werden (Figur 1c, d).

Die Auftrennung von HtH1 und HtH2 wurde gemäß des Verfahrens zur selektiven Dissoziation gemäß Harris et al., 1995, supra, durchgeführt. In Ammoniummolybdat/Polyethylenglykol war HtH1 im Oligomeren-Zustand (Di-Dekamere) vollständig stabil, wohingegen HtH2 vollständig in die Untereinheiten dissoziierte (Figur 1e). Dies ermöglichte die quantitative Sedimentation von HtH1 in der Ultrazentrifuge, wohin gegen der größte Teil des HtH2 im Überstand verblieb. Aus dem wieder aufgelösten Pellet wurden große Mengen an HtH1 durch Gelfiltrationschromatographie zur Homogenität aufgereinigt, welche auch geringe Mengen an reinem HtH2 ergab (Figur 1f). Die Fraktionen wurden durch native PAGE (Figur 1g) und gekreuzte IE (Figur 1h, i) untersucht. Das Verfahren der selektiven Dissoziation von HtH2 entfernte sämtliche Tri-Dekamere aus den Proben, welches nahelegt, daß die letzteren aus HtH2, jedoch nicht aus HtH1 aufgebaut sind (Figur 1e). Das selektive Dissoziationsverhalten von HtH2 und auch die Fähigkeit, Aggregate zu bilden, die größer als *in vivo* -Di-Dekamere sind, entspricht den Eigenschaften von KLH2. Umgekehrt ähnelt die Stabilität von HtH1 unter diesen Bedingungen und seine Unfähigkeit, sich zu größeren Aggregaten als Di-Dekameren zu assemblieren,

dem Verhalten von KLH1. Diese Verwandtschaft wird weiter belegt durch die Reaktion von Anti-KLH1 bzw. Anti-KLH2 Antikörpern gegen die beiden HtH-Typen (Figur 1j-m).

Beispiel 2:

Analyse der Organisation der HtH1-Untereinheit

Die acht funktionellen Einheiten (FU's, häufig als "funktionelle Domänen" bezeichnet), die eine Mollusken-Hämocyanin-Untereinheit bilden, unterscheiden sich in der Primärstruktur und weisen keine immunologische Kreuz-Reaktivität auf, wie sich aus gekreuzter IE ergibt. Im Fall der aufgereinigten HtH1 Untereinheit (Figur 1g, h) wurden geringe Konzentrationen fünf unterschiedlicher Proteasen (Elastase, V8-Protease, Papain, Trypsin und Chymotrypsin) verwendet, welche die Peptid-Bindungen zwischen benachbarten FU's von KLH1 und KLH2 gespalten hatten (Gebauer et al., 1994, supra, Söhnngen et al., 1997, supra). Die Spaltprodukte wurden durch gekreuzte IE und SDS-PAGE (Fig. 2) untersucht. Elastase-Behandlung erzeugt acht einzelne FU's, abgeleitet aus der Anzahl an unterschiedlichen Immunpräzipitations-Gipfeln in der gekreuzten IE (Fig. 2a) und mit dem scheinbaren Molekulargewicht von ca. 50 kDa des Hauptteils der Spaltprodukte in SDS-PAGE (Fig. 2b). Ein weiterer Präzipitationsgipfel wurde als FU-Dimer erkannt, welches durch unvollständige Spaltung des Segments ab (Fig. 2a) entstand.

Durch HPLC-Verfahren mit einer Mono-Q-Säule (Fig. 3a) wurden zwei der Elastase-

Spaltprodukte in einer ausreichenden Reinheit erhalten, um durch "crossed-line-IE" ihre klare Zuordnung zu zwei der acht Präzipitationsgipfel (Fig. 2c,d) zu ermöglichen. Die anderen vier Proteasen wiesen unterschiedliche Spaltmuster auf, die aus Gemischen einzelner FU's und größerer Fragmente, enthaltend zwei, drei oder mehrere FU's bestanden (z.B. Fig. 2 e,f). Viele von ihnen waren durch das HPLC-Verfahren in ausreichender Menge angereichert (Fig. 3b-e), um ihre Identifikation in ihren entsprechenden SDS-PAGE- und gekreuzten IE-Mustern zu ermöglichen. Eine Anzahl dieser Komponenten wurde N-terminal durch Blot-Transfer von SDS-Gelen auf ProBlot®-Membrane (Tabelle I) sequenziert. Die Resultate wurden mit N-terminalen Sequenzen verglichen, die aus dem scheinbar orthologen Protein in *Megathura crenulata*, KLH1 (Tabelle I), erhalten worden waren, von dem die vollständige FU-Anordnung verfügbar ist (Söhnngen et al., 1997, supra; vgl. Fig. 5b). Das Ergebnis des gesamten Ansatzes führte zur Bestimmung der vollständigen FU-Anordnung innerhalb der HtH1-Untereinheit (Fig. 2a).

Insbesondere ergab die Spaltung der HtH1-Untereinheit (1-abcdefgh) mit V8-Protease vier Präzipitationsgipfel in der gekreuzten IE (Fig. 2e). Das SDS-PAGE-Verfahren zeigte fünf unterschiedliche Fragmente (Fig. 2f): 220 kDa (5 FU's), 185 kDa (4 FU's), 100 kDa (2 FU's), 55 kDa (1 FU) und 46 kDa (1 FU). Das 100 kDa-Fragment wurde durch das HPLC-Verfahren (Fig. 3b) isoliert und durch N-terminale Sequenzierung als 1-ab identifiziert, da die Sequenz identisch zu derjenigen der intakten Untereinheit war (Tabelle I). In dem "crossed-line" IE-Verfahren verschmolz 1-ab mit drei Präzipitationsgipfeln des Elastase-Spaltmusters. Aufgrund der Auswertung ergab sich, daß sie die Fragmente 1-ab, 1-a bzw. 1-b darstellen (Fig. 2g). Jedoch verblieb es unklar, welcher Gipfel 1-a und welcher 1-b darstellt. In einem zweiten Schritt wurde das HPLC-aufgeklärte 1-ab durch Elastase in seine Komponenten-FU's gespalten, aus denen eine durch das native PAGE-Gel-Streifenverfahren eluiert werden konnte und dem Elastase-Muster durch das "crossed-line" IE-Verfahren (Fig. 2h) zugeordnet und N-terminal sequenziert wurde. Diese Komponente wies die gleiche N-terminale Sequenz wie die gesamte Untereinheit auf und war deshalb identisch zu 1-a. Somit ist die zweite FU des 100 kDa-Fragments 1-b (Fig. 2a; Tabelle I). Ebenso wurden HPLC-aufgeklärtes 1-c und 1-h erhalten (Fig. 3b), durch N-terminale-Sequenzähnlichkeiten zu den entsprechenden FU's in KLH1 (Tabelle I) identifiziert und durch das "crossed-line" IE-Verfahren ihren entsprechenden Präzipitationsgipfeln in dem Elastase-Muster zugeordnet (Fig. 2i,j). Weiterhin wurden 1-a, 1-b, 1-c und 1-h identifiziert (Fig. 2a). Unter Verwendung von Papain zur Untereinheiten-Spaltung wurden fünf unterschiedliche Gipfel in dem gekreuzten IE-Verfahren (Fig. k) erhalten. Aus einer solchen Probe wurde ein 100 kDa-Fragment (2 FU's) durch das HPLC-Verfahren (Fig. 3c) aufgereinigt, welches gemäß des "crossed-line" IE-Verfahrens das bereits identifizierte FU 1-h und eine der vier noch nicht identifizierten FU's enthielt und deshalb 1-gh darstellen muß (Figs. 2k, 3c). Tatsächlich wies dieses Fragment eine N-terminale Sequenz auf, die Ähnlichkeiten zu KLH1-g (Tabelle I) zeigte. Zur weiteren Bestätigung wurde das HPLC-aufgeklärte Fragment 1-gh mit Elastase in seine Bestandteil-FU's gespalten, aus denen 1-g aufgereinigt und durch N-terminale Sequenzierung identifiziert wurde. Es wurde durch "crossed-line" IE-Verfahren seinem Gipfel in dem Elastase-Spaltmuster zugeordnet (Fig. 2l).

Das 220 kDa-Fragment aus der V8-Protease-Spaltung (Fig. 2e, f) wurde HPLC-aufgereinigt (Fig. 3b) und fusionierte im "crossed-line" IE-Verfahren mit 1-h, 1-g und drei noch nicht identifizierten Gipfeln des Elastase-Spaltmusters. Weiterhin wurde das 185 kDa-Fragment in ausreichender Reinheit erhalten (Fig. 2e, f; 3b), und es wurde gezeigt, daß sie die gleichen Komponenten mit Ausnahme von 1-h enthielten. Dies legte nahe, daß das 22 kDa und das 185 kDa-Fragment 1-defgh bzw. 1-defg sind. Tatsächlich war die N-terminale Sequenz praktisch identisch und zeigte weiterhin Ähnlichkeit zu KLH1-d (Tabelle I). Die Spaltung der HtH1-Untereinheit mit Trypsin ergab eine Vielzahl an Komponenten in dem Molekulargewichtsbereich von ein bzw. zwei FU's (Fig. 2m). Mehrere der Komponenten wurden in HPLC-Fractionen angereichert (Fig. 3d). Ein 100 kDa-Fragment erwies sich als besonders nützlich, da es die gleich N-terminale Sequenz wie das Fragment 1-defg aus der V8-Protease-Spaltung aufwies (Tabelle I); deshalb sollte das 100 kDa-Fragment 1-de darstellen. In dem "crossed-lined" IE-Verfahren fusionierte diese Komponente mit zwei der drei noch nicht identifizierten FU-Gipfeln des Elastase-Spaltmusters (Fig. 2n), welches deshalb 1-d und 1-e darstellen sollte, und somit eine einzige Möglichkeit für 1-f übrig ließ. Das "crossed-line" IE-Verfahren zeigte auch, daß in der 1-de Fraktion weiterhin FU 1-f vorhanden war (Fig. 2n). Die Identifikation von 1-f bestätigte durch Spaltung der Untereinheit mit Chymotrypsin (Fig. 2o) und nachfolgendem HPLC-Verfahren (Fig. 3e). Diese Spaltung ergab unter anderem ein 95 kDa-Fragment (2 FU's) welches im "crossed-line" IE-Verfahren mit 1-g und einem zweiten Gipfel (Fig. 2p) fusionierte und deshalb entweder 1-gh (welches ausgeschlossen werden konnte, da 1-h bereits identifiziert war), oder 1-fg (welches sinnvoll erscheint aufgrund des weiteren betreffenden Peaks, der zu dem verbleibenden Kandidaten identisch war) sein könnte. Tatsächlich zeigte dieses Fragment eine neue N-terminale Sequenz, welche in gewisser Weise zu KLH1-f ähnlich ist (Tabelle I). Das letzte Problem bestand nun darin, die zwei verbleibenden FU-Gipfel 1-d bzw. 1-e zuzuordnen. Dies wurde unter Verwendung von HPLC-isolierten FU's aus Proben gelöst, in denen die Untereinheit mit Elastase gespalten worden war. (Fig. 2c, d; 3a). Die saurere Komponente in dem gekreuzten IE-Verfahren war 1-d abgeleitet aus seiner N-terminalen Sequenz, welche identisch ist mit der von 1-defgh (Fig. 2c; Tabelle I), wohingegen die basischere Komponente des 1-d/1-g-Paares eine neue N-terminale Sequenz (Tabelle I) aufwies und des-

halb 1-e (Fig. 2a) sein mußte. Somit war die Struktur der funktionellen Einheiten der Untereinheit HtH1 aufgeklärt.

Beispiel 3:

Vergleich der Molekulargewichte und N-terminalen Sequenzen der biochemisch isolierten funktionellen Einheiten (FU's) aus HtH1 und KLH1. Die unterschiedlichen FU's, jede mit einer intakten binukleären Kupfer-Bindungsstelle, wurden durch limitierte Proteolyse als globuläre Segmente aus ihrer größeren Einheit freigesetzt; vgl. Abschnitt "Isolierung und Analyse der Einheiten aus HtH1". Die KLH1-Daten wurden aus Söhngen et al., supra, entnommen. Die Zuordnung als tatsächliche Einheit erfolgte aufgrund des Molekulargewichtes und des immunologischen Verhaltens (vgl. Fig. 2). Das ungewöhnlich niedrige Molekulargewicht von isoliertem HtH1-d könnte bedeuten, daß ein großes Peptid C-terminal abgespalten wurde.

TABELLE I

Funktionelle Einheit	Masse (kDa)	N-terminale Sequenz
HtH1-a	53	DNVVRKDVSHLTDDEVQ
KLH1-a	50	ENLVVRKDVERL
HtH1-b	48	?
KLH1-b	45	?
HtH1-c	46	FEDEKHSLRIRKNVDSLTPPEENTNERLR

TABELLE I

Funktionelle Einheit	Masse (kDa)	N-terminale Sequenz
KLH1-c	45	KVPRSRRLIRKNVDRLTPSE
HtH1-d	40	VEEVTGASHIRKNLNDLNTGEM
KLH1-d	50	EVTSANRIRKNIENLS
HtH1-e	49	ILDHDHEEEILVRKNIIDLSP
KLH1-e	50	?

HtH1-f	50	KLNSRKHTPNR VR HELSSLSSRDIASLKA
KLH1-f	45	HHLSXNK VR HDLSTL
HtH1-g	45	DHQSGSIAGSG VR KDVNTLTKAETDNLRE
KLH1-g	45	SSMAGHF VR KDINTLTP
HtH1-h	55	DEHHDDRLADV LIR KEVDFLSLQEANA IKD
KLH1-h	60	HEDHHEDIL VR KNIHSL

Beispiel 4:

Klonierung von Hämocyanin-cDNA

1. Zur Klonierung der cDNA von Hämocyanin wurde mRNA aus dem Mantelgewebe des jeweiligen Mollusken isoliert. Der erste cDNA-Strang wurde durch reverse Transkription mit Oligo(dT) als Primer erhalten. Der zweite Strang wurde durch konventionelle Synthese mit random Primern erhalten. Die so erhaltene cDNA wurde in einen Lambda-Expressionsvektor kloniert unter Bildung einer cDNA-Expressionsbibliothek. Unter Verwendung eines anti-Hämocyanin-Antikörpers wurde die Bibliothek unter geeigneten Bedingungen abgesucht, wobei positive Klone erhalten wurden. Diese positiven Klone wurden isoliert, sequenziert und charakterisiert.

2. Aus dem N-terminalen Bereich eines erhaltenen positiven Klon wurde eine cDNA-Sonde hergestellt, mit der die cDNA-Bibliothek abgesucht wurde. Die erhaltenen positiven Klone werden wiederum isoliert, sequenziert und charakterisiert.
3. Um noch weiter 5' gelegene Sequenzen zu erhalten, wurde eine weitere Expressionsbibliothek aus cDNA hergestellt, die mit Hilfe einer Kombination von Hämocyaninspezifischen und „random“-Primern erhalten wurde. Diese cDNA-Bibliothek wurde mit cDNA-Sonden, die den „N-terminalen“ Bereichen der unter (2.) erhaltenen positiven Klone entsprechen, abgesucht. Die erhaltenen positiven Klone wurden isoliert, sequenziert und charakterisiert.

Beispiel 5:**Klonierung von Hämocyanin-Genen**

Genomische DNA wurde gemäß Standardverfahren isoliert. Die PCR-Reaktion wurde mit Hilfe von Hämocyanin-spezifischen Primern durchgeführt, um die Genabschnitte der interessierenden Hämocyanine zu amplifizieren. Die erhaltenen Amplifikationsprodukte wurden in einen geeigneten Vektor (beispielsweise pGem T oder pGem T easy (Promega, Mannheim) kloniert, sequenziert und charakterisiert.

Beispiel 6:**Rekombinante Expression von Hämocyanin**

Mit einem cDNA-Klon, der die kodierende Sequenz für HtH-1d enthält, wurde eine PCR-Reaktion durchgeführt, um spezifisch die kodierende Sequenz der Domäne 1d zu amplifizieren. Als Primer wurden synthetisch hergestellte Oligonukleotide verwendet. Primer 1 (stromaufwärts) umfaßt sechs Nukleotide des Endes der Domäne HtH-1c, eine *SacI*-Schnittstelle und 12 Nukleotide des Endes der Domäne HtH-1d. Primer 2 (stromabwärts) umfaßt sechs Nukleotide des Anfangs der Domäne HtH1-e, eine *SalI*-Schnittstelle und eine HtH1-d spezifische Sequenz.

PCR-Bedingungen:

2	min	95°C
30	sec	95°C
30	sec	55°C
1	min	72°C
35	Zyklen	
10	min	72°C

Das Amplifikat wurde in dem pGEM T easy PCR Klonierungsvektor (Promega) in XL-1 Blue (Stratagene) kloniert. Nach Isolation des rekombinanten Plasmids und Restriktion mit *SacI* und *SalI* konnte die cDNA der Domäne 1d isoliert werden. Der Expressionsvektor pQE30 (Qiagen) wurde ebenfalls mit den entsprechenden Enzymen restringiert.

Anschließend wurde die Ligation zwischen der Hth-1d-cDNA (restringiert mit *SacI* und *Sall*) und pQE (restringiert mit *SacI* und *Sall*) durchgeführt. Somit ist eine gerichtete Klonierung der cDNA, kodierend für Hth-1d, in einen Expressionsvektor möglich. Die Expression von Hth1-d in pQE in XL-1 Blue erfolgt gemäß Herstelleranleitung. Die Expression weiterer Hth1-, Hth2- oder KLH1- oder KLH2-Domänen kann analog erfolgen.

SEQUENZPROTOKOLL

<110> Biosyn Arzneimittel GmbH

<120> Nukleinsäuremolekül, umfassend eine für ein Hämocyanin kodierende Nukleinsäuresequenz

<130> P30370B-01996

<140>

<141>

<160> 48

<170> PatentIn Ver. 2.1

<210> 1

<211> 1269

<212> DNA

<213> Haliotis tuberculata

<220>

<223> HtH1-a

<400> 1

```

ggcttggttca gtttctactc gtcgcccttg tggcgggggc tggagcagac aacgtcgtca 60
gaaaggacgt gagtcacctc acggatgacg aggtgcaagc tctccacggc gccctccatg 120
acgtcactgc atctacaggg cctctgagtt tcgaagacat aacatcttac catgccgcac 180
cagcgtcgtg tgactacaag ggacggaaga tcgcctgctg tgtccacggc atgcccagtt 240
tccccttctg gcacagggca tatgtcgtcc aagccgagcg ggcaactgtt tccaaacgga 300
agactgtcgg aatgccttac tgggactgga cgcaaacgct gactcactta ccatctcttg 360
tgactgaacc catctacatt gacagtaaag gtggaaaggc tcaaaccaac tactggtacc 420
gcggcgagat agcgttcatc aataagaaga ctgcgcgagc tgtagatgat cgcctattcg 480
agaagggtgga gcctggtcac tacacacatc ttatggagac tgtcctcgac gctctcgaac 540
aggacgaatt ctgtaaattt gaaatccagt tcgagttggc tcataatgct atccattact 600
tggttggcgg taaatttgaa tattcaatgt caaacttgga atacacctcc tacgacccca 660
tcttcttctt ccaaccattc aacgttgacc gctcttctgc catctggcag cgtcttcagg 720
aactgcgagg aaagaatccc aatgcaatgg actgtgcaca tgaactcgct caccagcaac 780
tccaaccctt caacagggac agcaatccag tccagctcac aaaggaccac tcgacacctg 840
ctgacctctt tgattacaaa caacttggat acagctacga cagcttaaac ctgaatggaa 900
tgacgccaga acagctgaaa acagaactag acgaacgcca ctcaaagaa cgtgcgtttg 960
caagcttccg actcagtggc tttgggggtt ctgccaacgt tgttgctctat gcatgtgtcc 1020
ctgatgatga tccacgcagt gatgactact gcgagaaaagc aggcgacttc ttcatctctg 1080
ggggtcaaag cgaaatgccg tggagattct acagaccctt cttctatgat gtaactgaag 1140
cggtacatca ccttgagtc cgcctaagtg gccactacta tgtgaaaaca gaactcttca 1200
gcgtgaatgg cacagcactt tcacctgatc ttcttctca accaactgtt gcctaccgac 1260
ctgggaaaag

```

<210> 2

<211> 1257

<212> DNA

<213> Haliotis tuberculata

<220>

<223> HtH1-b

<400> 2

```

gtcaccttga cccacctgtg catcatcgcc acgatgacga tcttattgtt cgaaaaaata 60
tagatcattt gactcgtgaa gaggaatacg agctaaggat ggctctggag agattccagg 120
ccgacacatc cgttgatggg taccaggcta cagtagagta ccatggcctt cctgctcgtt 180
gtccacgacc agatgcaaaa gtcaggttcg cctggtgtat gcatggcatg gcatccttcc 240
ctcactggca ccggctgttc gttaccaggg tgggaagatgc tcttgtagcg cgtggatcgc 300

```

```

ctatcgggtgt tccttacttg gactggacaa aacctatgac tcaccttcca gacttggcat 360
caaatgagac gtacgtagac ccgtatggac atacacatca taatccattc ttcaatgcaa 420
atatatcttt tgaggaggga caccatcaca cgagcaggat gatagattcg aaactgtttg 480
ccccagtcgc ttttggggag cattcccatc tgtttgatgg aatcctgtac gcatttgagc 540
aggaagattt ctgcgacttt gagattcagt ttgagttagt ccataattct attcatgctg 600
ggataggcgg ttccgaagat tactccatgg ccaccctgca ttacacagcc tttgaccca 660
ttttctacct tcatcattcc aatgtcgatc gtctatgggc aatctggcaa gctcttcaaa 720
tcaggagaca caagccatat caagcccact gtgcacagtc tgtggaacag ttgccaatga 780
agccatttgc tttcccatca cctcttaaca acaacgagaa gacacatagt cattcagtc 840
cgactgacat ttatgactac gaggaagtgc tgcactacag ctacgatgat ctaacgtttg 900
gtgggatgaa ccttgaagaa atagaagaag ctatacatct cagacaacag catgaacgag 960
tcttcgctggg atttctcctt gctggaatag gaacatctgc acttggtgac attttcataa 1020
ataaaccggg gaaccaacca ctcaaagctg gagatattgc cattcttggt ggtgccaaag 1080
aaatgccttg ggcgtttgac cgcttgatata aggtcgaaat aactgactca ttgaagacac 1140
tttctctcga tgtcgatgga gattatgaag tcacttttaa aattcatgat atgcacggaa 1200
acgctcttga tacggacctg attccacacg cagcagttgt ttctgagcca gctcacc 1257

```

<210> 3

<211> 1242

<212> DNA

<213> *Haliotis tuberculata*

<220>

<223> HtH1-c

<400> 3

```

ctacctttga ggatgaaaag cacagcttac gaatcagaaa aaatgtcgac agcttgactc 60
ctgaagaaac aaatgaactg cgtaaagccc tggagcttct tgaaaatgat catactgcag 120
gtggattcaa tcagcttggc gccttccatg gagagcctaa atggtgccct aatcctgaag 180
cggagcacia gggtgcatgc tgtgttcatg gcatggctgt tttccctcat tggcacaggc 240
ttcttgctct ccaggcggag aatgctctta gaaagcatgg gtacagtggg gctctaccat 300
actgggattg gactcgcccc ctttcccaac ttctgtatct ggtagtcat gaggagtata 360
cagatccttc cgaccatcac gtgaagcata acccgtggtt caatggccac atcgatacag 420
taaatcagga taccaccaga agcgtagcgg aggatcttta tcaacaacct gaatttggac 480
atttcacgga tattgctcaa caagtctctt tagcattaga acaagatgac ttctgttcgt 540
ttgaagtgca gtatgagatt tcccataatt ttatccatgc acttgtagga ggaaccgacg 600
cttatggcat ggcacgctg agatatacag catacgatcc aatctttttc ttgcatcatt 660
caaacaccga caggatctgg gctatttggc aatccctgca aaaatacaga ggcaaaccgt 720
acaacactgc caactgcgcc atagaatcta tgagaaggcc cctgcaacca tttggactaa 780
gcagtgccat taaccctgac agaatcacca gagagcatgc tatcccggtt gatgtcttca 840
actatagaga taaccttcat tacgtatatg ataccctgga atttaattgt ttgtcgattt 900
cacaacttga tagagagctg gaaaaaatca agagtacaga aagagtattt gctggattct 960
tgctgtcggg gattaaaaaa tctgtctctt tgaaattcga agtttgtact ccacttgata 1020
attgtcataa agcaggggag ttttatctac tcggggacga aaacgagatg gcttgggcct 1080
atgaccgact tttcaagtat gatattactc aggttctgga agcaaaccat ctacacttct 1140
atgatcatct cttcattcgc tacgaagtct ttgatcttaa aggagtgagt ttgggaactg 1200
acctgttcca cactgcaaat gtggtacatg attccggcac ag 1242

```

<210> 4

<211> 1239

<212> DNA

<213> *Haliotis tuberculata*

<220>

<223> HtH1-c

<400> 4

```

gcacccgtga tcgtgataac tacgttgaag aagttactgg ggccagtcac atcaggaaga 60
atttgaacga cctcaatacc ggagaaatgg aaagccttag agctgctttc ctgcatattc 120
aggacgacgg aacatatgaa tctattgccc agtaccatgg caaaccaggc aaatgtcaat 180
tgaatgatca taatattgcg tgttgtgtcc atggtatgcc taccttcccc cagtggcaca 240
gactgtatgc gtttcaggtg gagaatgctc tcctaaacag gggatctggg gtggctgttc 300
cttactggga gtggactgct cccatagacc atctacctca tttcattgat gatgcaacat 360
acttcaattc ccgacaacag cggtagcacc ctaacccttt cttcagggga aaggttactt 420
ttgaaaacgc agtcacaaca agggacccac aagccgggct cttcaactca gattatatgt 480
atgagaatgt tttacttgca ctggagcagg aaaattattg tgactttgaa attcagtttg 540

```

agcttggttca	taacgcactt	cattccatgc	tgggaggttaa	agggcagtag	tccatgtcct	600
ccctggacta	ttctgcgttt	gatcccgtct	tcttcctaca	tcatgccaac	acggacagac	660
tgtgggcaat	ctggcaggaa	ctacaaagat	tccgagaact	gccttatgaa	gaagcgaact	720
gtgcaatcaa	cctcatgcat	caaccactga	agccgttcag	tgatccacat	gagaatcacg	780
acaatgtcac	tttgaaatac	tcaaaaccac	aggacggatt	cgactaccag	aaccacttcg	840
gatacaagta	tgacaacott	gagttccatc	acttatctat	cccaagtctt	gatgctaccc	900
tgaagcaaag	gagaaatcac	gacagagtgt	ttgcgggctt	ccttcttcat	aacataggaa	960
cttctgctga	cataactatc	tacatatgtc	tgccctgacg	acggcgtggc	aatgactgca	1020
gtcatgaggc	gggaacattc	tatatcctcg	gaggcgaaac	agagatgcct	tttatctttg	1080
accgtttgta	taaatttgaa	atcaccaaac	cactgcaaca	gttaggagtc	aagctgcatg	1140
gtggagtttt	cgaactggag	cttgagatca	aggcatacaa	cggttcctat	ctggatcccc	1200
atacctttga	tccaactatc	atctttgaac	ctggaacag			1239

<210> 5

<211> 1260

<212> DNA

<213> *Haliotis tuberculata*

<220>

<223> HtH1-e

<400> 5

tacccatata	cttggaccac	gacccatgagg	aagagatact	tgtaggaag	aatataattg	60
ttttgagccc	aagggagagg	gtttctctag	tcaaagcttt	gcaaagaatg	aagaatgac	120
gtctccgtga	tgggtaccaa	gccattgcct	ctttccatgc	cctgccacca	ctctgtccca	180
atccatctgc	agctcaccgt	tatgcttgct	gtgtccatgg	catggctaca	tttccccagt	240
ggcacagact	gtacactgtt	caggttcagg	atgccctgag	gagacatggg	tcacttggtg	300
gtattcctta	ctgggactgg	acaaaaccag	tcaacgagtt	acccgagctt	ctttcttcag	360
caacatttta	tcatccaatc	cggaatatta	atattttcaa	tccattcctc	ggggctgaca	420
tagaattttga	aggaccgggc	gttcatacac	agaggcacat	aaatactgag	cgccgtgttc	480
acagtgggga	tcatgacgga	taccacaact	ggttcttcga	aactgtttct	tttgctttgg	540
aacaggaaga	ttactgcat	tttgaaatac	aatttgagat	agcccataat	ggcatccaca	600
catggattgg	tggaaagcga	gtatatggca	tgggacacct	tcactatgca	tcatatgac	660
caatttttcta	catccaccat	tcacagacgg	acagaatatg	ggctatttgg	caagagctgc	720
agaagtacag	gggtctatct	ggttcggaag	caaactgtgc	cattgaacat	atgagaacac	780
ccttgaagcc	tttcagcttt	gggccaccct	acaatttgaa	tagtcatacg	caagaatatt	840
caaagcctga	ggacacgttt	gactataaga	agtttgata	cagatatgat	agtctggaat	900
tggagggggcg	atcaatttct	cgcatgtgat	aacttatcca	gcagagacag	gagaaagaca	960
gaacttttgc	agggttcctc	cttaaaggtt	ttggtacatc	cgcatctgtg	tcattgcaag	1020
tttgacagagt	tgatcacacc	tgtaaagatg	cgggctattt	cactattctg	ggaggatcag	1080
ccgaaatgcc	atgggcattc	gacaggtctt	ataagtatga	cattactaaa	actcttcacg	1140
acatgaacct	gaggcacgag	gacactttct	ctatagacgt	aactatcacg	tcttacaatg	1200
gaacagtact	ctcgggagac	ctcattcaga	cgccctccat	tatatttgta	cctggagcgc	1260

<210> 6

<211> 1251

<212> DNA

<213> *Haliotis tuberculata*

<220>

<223> HtH1-f

<400> 6

ataaactcaa	ctcacggaaa	catacaccta	acagagtccg	ccatgagcta	agtagcctta	60
gttcccgtga	catagcaagc	ttgaaggcag	ctttgacaag	ccttcaacat	gataatggga	120
ctgatggtta	tcaagctatt	gctgccttcc	atggcggttcc	tgcgcagtag	cacgagccat	180
ctggacgtga	gatcgccgtg	tgcatccacg	gcatggcgac	gtttcctcac	tggcaccggt	240
tgtacactct	gcagttggag	caagcgctgc	gcagacacgg	gtccagtgtt	gctgttccat	300
actgggactg	gaccaagcca	atcaccgaac	tgccacacat	tctgacagac	ggagaatatt	360
atgacgtttg	gcaaaatgcc	gtcttggcca	atccgtttgc	aagaggttat	gtgaaaatta	420
aagatgcatt	tacgggtgaga	aatgtccagg	aaagtctgtt	caaaatgtca	agttttggaa	480
agcactcgct	tctgtttgac	caggctttgt	tggctcttga	acaaactgac	tactgtgact	540
tcgaagtcca	gtttgaagtg	atgcataaca	cgatccatta	tctcgtagga	gggcgtcaaa	600
cgtacgcctt	ctcctctctc	gagtattcct	catacgatcc	aatcttcttt	attccaccat	660
cgtttgattg	caaaatatgg	gctgtatggc	aagaactgca	aagcaggaga	catctacagt	720
ttagaacagc	tgattgtgct	gtgggcctca	tgggtcaggc	aatgaggcct	ttcaacaagg	780

```

atttcaacca caacttc accaagaagc acgcagtcct taataagta tttgattatg 840
aagatcttgg ctataactat gacaaccttg aaatcagtggt tttaaactta aatgagatcg 900
aggcggttaat agcaaaacgc aagtcacatg cttagagtcct tgctgggttc ctgttgtttg 960
gattaggaac ttcggctgat atacatctgg aaatttgcaa gacatcgga aactgccatg 1020
atgctgggtgt gattttcatc cttggagggt ctgcagagat gcattgggca tacaaccgcc 1080
tctacaagta tgacattaca gaagcattgc aggaatttga catcaaccct gaagatgttt 1140
tccatgctga tgaaccattt ttcctgaggc tgcgggtgtg tgctgtgaat ggaactgtca 1200
ttccatcgtc tcatcttcac cagccaacga taatctatga accaggcgaa g 1251

```

<210> 7

<211> 1209

<212> DNA

<213> *Haliotis tuberculata*

<220>

<223> HtH1-g

<400> 7

```

atcaccatga cgaccatcag tcgggaagca tagcaggatc cggggtccgc aaggacgtga 60
acaccttgac taaggctgag accgacaacc tgaggagggc gctgtggggt gtcattggcag 120
accacgggtcc caatggcttt caagctattg ctgctttcca tggaaaacca gctttgtgtc 180
ccatgcctga tggccacaac tactcatggt gtactcacgg catggtctacc ttcccacact 240
gcatcgctct ctacaccaag cagatggagg atgcaatgag ggcgcatggg tctcatgtcg 300
cctgcctcta ctgggactgg actgtgcctt tcaccacact gccaacactg gtcaccgaca 360
gggacaacaa ccccttccaa catggacaca ttgattatct caatgtcagc acaactcgat 420
ctcccggaga catgctgttc aacgaccccg agcatggatc agagtcgttc ttctacagac 480
aagtcctctt agctctggaa caaactgatt tctgcaaatt cgaagtccag tttgagataa 540
cccacaatgc catccattcc tggacagggt gccacagccc ctacggaatg tccactctcg 600
acttcaactgc ctacgatcct ctcttctggc ttccaccact caacaccgac agaactctggg 660
ctgtctggga agctttgcaa gaatacagag gacttccata caaccatgcc aattgtgaga 720
tccaggcaat gaaaacgccc ctgaggcctt tcagtgacga tatcaaccac aaccagtcga 780
caaaggctaa cggaagcca ttagatgtgt tcgagtataa tcggttgagc ttccagtacg 840
acaacctcat cttccatgga tacagtattc cggaacttga tcgctgtgctt gaagaaagaa 900
aggaggagga cagaatattt gctgccttcc ttctcagtggt aatcaagcgt agtgctgatg 960
tagtgttcga catatgccag ccagaacacg aatgtgtgtt cgcagggact tttgcgattt 1020
tgggagggga gctagaaatg ccctgtctct tcgacagact gttccgctat gatataacca 1080
aggtgatgaa cgagctacac ctgaggcatg actctgactt taccttcagg gtgaagattg 1140
tcggcaccga cgaccacgag cttccttcag acagtgtcaa agcaccaact attgaatttg 1200
aaccgggagc 1251

```

<210> 8

<211> 1535

<212> DNA

<213> *Haliotis tuberculata*

<220>

<223> HtH1-h

<400> 8

```

tgcacagagg cggaaccac gaagatgaac accatgatga cagactcgca gatgtcctga 60
tcaggaaaga agttgacttc ctctccctgc aagaggccaa cgcaattaag gatgcactgt 120
acaagctcca gaatgacgac agtaaagggg gctttgaggc catagctggc tatcacgggt 180
atcctaatat gtgtccagaa agaggtaacc acaagtatcc ctgctgtgtc cacggaatgc 240
ccgtgttccc ccaactggcag cgctgcata ccattcagat ggagagagct ctgaaaaacc 300
atggctctcc aatgggcatt ccttactggg attggacaaa gaagatgtcg agtcttccat 360
ctttcttttg agattccagc aacaacaacc ctttctacaa atattacatc cggggcgtgc 420
agcacgaaac aaccagggac attaatcaga gactctttaa tcaaaccaag tttggtgaat 480
ttgattacct atattacct actctgcaag tcttgaggga aaactcgtac tgtgactttg 540
aagttcagta tgagatcctc cataacgccc tccactcctg gcttggagga actggaaagt 600
attccatgtc taccctggag cattcggcct ttgacctgt cttcatgatt caccactcga 660
gtttggatag aatctggatc ctttggcaga agttgcaaaa gataagaatg aagccttact 720
acgcattgga ttgtgctggc gacagactta tgaaagacce cctgcacccc ttcaactacg 780
aaaccgttaa tgaagatgaa ttcacccgca tcaactcttt cccaagcata ctgtttgacc 840
actacaggtt caactatgaa tacgataaca tgagaatcag gggtcaggac atacatgaac 900
ttgaagaggt aattcaggaa ttaagaaaca aagatcgcat atttgctggg tttgttttgt 960
cgggcttacg gatatacagc acagtgaagc tattcattca ttcgaaaaac gatataagtc 1020

```

acgaagaata	tgcaggagaa	tttgcagttt	tgggaggtga	gaaggagatg	ccgtgggcat	1080
atgaaagaat	gctgaaattg	gacatctccg	atgctgtaca	caagcttcac	gtgaaagatg	1140
aagacatccg	gtttagagt	gttttactg	cctacaacgg	tgacgttggt	accaccaggc	1200
tgtctcagcg	attcatcgtc	caccgtccag	cccatgtggc	tcacgacatc	ttggtaatcc	1260
cagtaggtgc	gggccatgac	cttccgccta	aagtcgtagt	aaagagcggc	accaaagtcg	1320
agttttacacc	aatagattcg	tcggtgaaca	aagcaatggt	ggagctgggc	agctatactg	1380
ctatggctaa	atgcatcggt	ccccctttct	cttaccacgg	ctttgaactg	gacaaagtct	1440
acagcgtcga	tcacggagac	tactacattg	ctgcaggtag	ccacgcgttg	tgtgagcaga	1500
acctcaggct	ccacatccac	gtggaacacg	agtag			1535

<210> 9

<211> 1003

<212> DNA

<213> *Haliotis tuberculata*

<220>

<223> Hth2-b

<400> 9

cacagactgt	tcgtcaccca	ggtggaagat	gctctgatca	ggcgaggatc	gcctataggg	60
gtccctact	gggactggac	tcagcctatg	gcgcctctcc	caggacttgc	agacaacgcc	120
acctatagag	atcccatcag	cggggacagc	agacacaacc	ccttccacga	tggtgaagtt	180
ccctttgaaa	atggacgtac	agaacgtcac	ccagatagta	gattgtttga	acaaccttta	240
tttgcaaac	atacgctct	cttcgacagt	atagtctatg	cttttgagca	ggaggacttc	300
ggcgattttg	aagttcaatt	tgagatgacc	cataataata	ttcacgcctg	gattgggtggc	360
ggcgagaagt	attccatgtc	ttctctacac	tacacagcct	tcgaccctat	cttctacctt	420
cgctactcca	acactgaccg	gctctgggca	atttggcaag	cgttgcagat	acgaagaaac	480
aggccttaca	aggctcattg	tgcttggtct	gaggaacgcc	agcctctcaa	acctttcgcc	540
ttcagttccc	cactgaacaa	caacgaaaaa	acctacgaaa	actcggtgcc	caccaacggt	600
tacgactacg	aaggagtcct	tggtataact	tatgatgacc	tcaacttcgg	gggcatggac	660
ctgggtcagc	ttgaggaata	catccagagg	cagagacaga	gagacaggac	ctttgctggt	720
ttctttctgt	cacatattgg	tacatcagcg	aatgttgaaa	tcattataga	ccatgggact	780
cttcatacct	ccgtgggcac	gtttgctggt	cttggcggag	agaaggagat	gaaatgggga	840
tttgaccgtt	tgtacaaaata	tgagattaca	gatgaactga	ggcaacttaa	tctccgtgct	900
gatgatgttt	tcagcatctc	tggttaaagta	actgatgttg	atggcagtga	gctgtcctct	960
gaactcatcc	catctgctgc	tatcatcttc	gaacgaagcc	ata		1003

<210> 10

<211> 1251

<212> DNA

<213> *Haliotis tuberculata*

<220>

<223> Hth2-c

<400> 10

ttgaccatca	ggaccgcat	catgacacaa	tcattaggaa	aaatgttgat	aatcttacac	60
ccgaggaaat	taattctctg	aggcgggcaa	tggcagacct	tcaatcagac	aaaaccgccg	120
gtggattcca	gcaaattgct	gcttttcacg	gggaacccaa	atggtgcca	agtcccgatg	180
ctgagaagaa	gttctcctgc	tgtgtccatg	gaatggctgt	cttccctcac	tggcacagac	240
tcctgaccgt	gcaaggcgag	aatgccctga	gaaagcatgg	atgtctcgga	gctctccctt	300
actgggactg	gactcggccc	ctgtctcacc	tacctgattt	ggttttggtg	agtagcagaa	360
ctacaccgat	gccatattcc	accgtggaag	cccgaacccc	ctggtacagc	ggccatattg	420
atacagttgg	tgttgacaca	acaagaagcg	tccgtcaaga	actgtatgaa	gctcctggat	480
ttggccatta	tactggggtc	gctaagcaag	tgcttctggc	tttgagcag	gatgacttct	540
gtgattttga	agtcagttt	gagatagctc	acaatttcat	tcacgctctt	gtcggcgga	600
gcgagccata	tggtatggcg	tcactccggt	acactactta	tgatccaatt	ttctacctcc	660
atcattctaa	cactgacaga	ctctgggcta	tatggcaggc	tctacaaaag	tacaggggca	720
aaccttacaa	ttccgccaac	tgcgccattg	cttctatgag	aaaaccctta	caaccctttg	780
gtctgactga	tgagatcaac	ccggatgatg	agacaagaca	gcatgctgtt	cctttcagtg	840
tctttgatta	caagaacaac	ttcaattatg	aatatgacac	ccttgacttc	aacggactat	900
caatctccca	gctggaccgt	gaactgtcac	ggagaaagtc	tcatgacaga	gtatttgccg	960
gatttttgct	gcatgggtatt	cagcagtcctg	cactagttaa	attcctttgtc	tgcaaatcag	1020
atgatgactg	tgaccactat	gctggtgaat	tctacatcct	tggtgatgaa	gctgaaatgc	1080
catggggcta	tgatcgtctt	tacaaatatg	agatcactga	gcagctcaat	gccctggatc	1140
tacacatcgg	agatagattc	ttcatcagat	acgaagcgtt	tgatcttcat	ggtacaagtc	1200

ttggaagcaa catc cc aaaccttctg tcatacatga cgaagggca g

1251

<210> 11

<211> 1244

<212> DNA

<213> *Haliotis tuberculata*

<220>

<223> HtH2-d

<400> 11

```
gtcaccatca ggctgacgag tacgacgaag ttgtaactgc tgcaagccac atcagaaaga 60
atttaaaaga tctgtcaaag ggagaagtag agagcctaag gtctgccttc ctgcaacttc 120
agaacgacgg agtctatgag aatattgcc aagctccagg caagcctggg ttgtgtgatg 180
ataacgggtcg caaggttgcc tgttgtgtcc atggaatgcc caccttcccc cagtggcaca 240
ggctctatgt cctccagggt gagaatgctt tgctggagag aggatctgcc gtctctgtgc 300
catactggga ctggactgaa acatttacag agctgccatc tttgattgct gaggtacct 360
atttcaatgc ccgtcaacaa acgtttgacc ctaatccttt cttcagaggt aaaatcagtt 420
ttgagaatgc tgttacaaca cgtgatcccc agcctgagct gtacgttaac aggtactact 480
accaaaacgt catgttggtt tttgaacagg acaactactg cgacttcgag atacagtttg 540
agatggttca caatgttctc catgcttggc ttggtggaag agctacttat tctatttctt 600
ctcttgatta ttctgcattc gaccctgtgt ttttccttca ccatgcgaac acagatagat 660
gtggggccat ctggcaggag ctgcagaggt acaggaagaa gccatacaat gaagcggatt 720
tgccattaa cctaattgcg aaacctctac atcccttcga caacagtgat ctcaatcatg 780
atcctgtaac ctttaaatat tcaaaaccca ctgatggctt tgactaccag aacaactttg 840
gatacaagta tgacaacctt gagttcaatc atttcagtat tcccaggctt gaagaaatca 900
ttcgtattag acaacgtcaa gatcgtgtgt ttgcaggatt cctccttcac aacattggga 960
catccgcaac tgttgagata ttcgtctgtg tccctaccac cagcggtgag caaaactgtg 1020
aaaacaaagc cggaacattt gccgtactcg gaggagaaac agagatggcg tttcattttg 1080
acagactcta caggtttgac atcagtgaac cactgaggga cctcggcata cagctggaca 1140
gccatgactt tgacctcagc atcaagattc aaggagtaaa tggatcctac cttgatccac 1200
acatcctgcc agagccatcc ttgatttttg tgccctggtc aagt 1244
```

<210> 12

<211> 1255

<212> DNA

<213> *Haliotis tuberculata*

<220>

<223> HtH2-e

<400> 12

```
tctttcctgc gtctgatgg gcattcagat gacatccttg tgagaaaaga agtgaacagc 60
ctgacaacca gggagactgc atctctgac catgctctga aaagtatgca ggaagaccat 120
tcacctgacg ggttccaagc cattgcctct ttccatgctc tgccaccact ctgcccttca 180
ccatctgcag ctcaccgtta tgcttgctgt gtccacggca tggctacatt tccccagtg 240
cacagattgt aactgtaca gttccaggat gcaactgagga gacatggagc tacggtaggt 300
gtaccgtatt gggattggct gcgaccgcag tctcacctac cagagcttgt caccatggag 360
acataccatg atatttggag taacagagat tcccccaatc ctttctacca agccaatatt 420
gagtttgaag gagaaaacat tacaacagag agagaagtca ttgcagacaa actttttgtc 480
aaaggtggac acgtttttga taaactgggt cttcaacaa gccatcctag cgctgagcag 540
gaaaactact gtgactttga gattcagttt gaaattcttc acaacggcgt tcacacgtgg 600
gtcggaggca gtcgtaccta ctctatcgga catcttcatt acgcattcta cgacctctt 660
ttctaccttc accatttcca gacagaccgt atttgggcaa tctggcaaga actccaggaa 720
cagagagggc tctcgggtga tgaggctcac tgtgctctcg agcaaatgag agaaccattg 780
aagcctttca gcttcggcgc tccttataac tggaaatcagc tcacacagga tttctcccga 840
cccgaggaca ccttcgacta caggaagttt ggttatgaat atgacaattt agaattcctg 900
ggaatgtcag ttgctgaact ggatcaatac attattgaac atcaagaaaa tgatagagta 960
ttcgtctggg tcctgttgag tggattcggg ggttccgcag cagttaattt ccaggtttgt 1020
agagctgatt ccacatgtca ggatgctggg tacttcaccg ttcttgggtg cagtgtctgag 1080
atggcgtggg catttgacag gctttacaaa tatgacatta ctgaaactct ggagaaaatg 1140
caccttcgat atgatgatga cttcacaatc tctgtcagtc tgaccgcca caacggaact 1200
gtcctgagca gcagtcta atcccaacacc agtgtcatat tccagcgggg acatc 1255
```

<210> 13

<211> 1248

<212> DNA
 <213> *Haliotis tuberculata*

<220>
 <223> HtH2-f

<400> 13
 gtgacataaa taccaggagc atgtcaccga accgtgttcg ccgtgagctg agcgatctgt 60
 ctgcgaggga cctgtctagt ctcaagtctg ctctgcgaga cctacaggag gatgatggcc 120
 ccaacggata ccaggctctt gcagccttcc atgggctacc agcaggctgc catgatagcc 180
 ggggaaatga gatcgcatgt tgcattcacg ggatgcgac cttccccag tggcacagac 240
 tgtacaccct gcagttggag atggctctga ggagacatgg atcatctgtc gccatccct 300
 actgggactg gacaaagcct atctccgaac tcccctcgct cttcaccagc cctgagtatt 360
 atgacccatg gcatgatgct gtggtaaaca acccattctc caaaggtttt gtcaaatttg 420
 caaataccta cacagtaaga gaccacagg agatgctgtt ccagctttgt gaacatggag 480
 agtcaatcct ctatgagcaa actcttcttg ctcttgagca aaccgactac tgtgattttg 540
 aggtacagtt tgaggtcctc cataacgtga tccactacct tgttggtgga cgtcagacct 600
 acgcattgtc ttctctgcat tatgcctcct acgacccatt cttctttata caccattcct 660
 ttgtggataa gatgtgggta gtatggcaag ctcttcaaaa gaggaggaaa cttccataca 720
 agcgagctga ctgtgctgtc aacctaataga ctaaaccaat gaggccattt gactccgata 780
 tgaatcagaa cccattcaca aagatgcacg cagttcccaa cacactctat gactacgaga 840
 cactgtacta cagctacgat aatctcgaaa taggtggcag gaatctcgac cagcttcagg 900
 ttgaaattga cagaagcaga agccacgacg gcgtttttgc tggattcttg cttcgtggaa 960
 cggaaacttc tgctgatgtc aggtttttgga tttgtagaaa tgaaatgac tgccacaggg 1020
 gtggaataat tttcatctta ggtggagcca aggaaatgcc atggtcattt gacagaaact 1080
 tcaagtttga tatcacccat gtactcgaga atgctggcat tagcccagag gacgtgtttg 1140
 atgctgagga gccattttat atcaaggttg agatccatgc tgttaacaag accatgatac 1200
 cgtcgtctgt gatcccagcc ccaactatca tctattctcc tggggaag 1248

<210> 14
 <211> 1207
 <212> DNA
 <213> *Haliotis tuberculata*

<220>
 <223> HtH2-g

<400> 14
 gtcgcgtgctg tgacagtgcg cactctgccca acattgctgg ctctgggggtg aggaaggacg 60
 tcacgaccct cactgtgtct gagaccgaga acctaagaca ggctcttcaa ggtgtcatcg 120
 atgatactgg tcccaatggt taccagcaa tagcatcctt ccacggaggt cctccaatgt 180
 gcgagatgaa cggccgcaag gttgcctggt gtgctcacgg tatggcctcc ttcccacact 240
 ggcacagact gtatgtgaag cagatggaag atgccctggc tgaccacggg tcacatatcg 300
 gcatccctta ctgggactgg acaactgcct tcacagagtt acccgccctt gtcacagact 360
 ccgagaacaa tcccttccat gagggtcgca ttgatcatct cgggtgaacc acgtcacgtt 420
 cccccagaga catgctgttt aacgacccag agcaaggatc agagtcgttc ttctatagac 480
 aagtccctcct ggctttggag cagactgact actgccagtt cgaagtccag tttgagctga 540
 cccacaacgc cattcactcc tggacagggtg gacgtagccc ttacggaatg tcgaccctcg 600
 agttcacagc ctacgatcct ctcttctggc ttcaccactc caacaccgac agaattctggg 660
 ctgtctggca agcactgcag aaataccgag gactcccata caacgaagca cactgtgaaa 720
 tccaggttct gaaacagccc ttgaggccat tcaacgatga catcaaccac aatccaatca 780
 ccaagactaa tgccaggcct atcgattcat ttgattatga gaggtttaac tatcagtatg 840
 acacccttag cttccatggg aagagcatcc ctgaactgaa tgacctgctc gaggaaagaa 900
 aaagagaaga gagaacattt gctgccttcc ttcttcgtgg aatcggttgc agtgctgatg 960
 tcgtctttga catctgccgg cccaatgggt actgtgtctt tgcaggaacc tttgctgtgc 1020
 tgggagggga gtagaaaatg ccttggtcct tcgacagact gttccgctat gacatcacca 1080
 gagtcatgaa tcagctccat ctccagtatg attcagattt cagtttcagg gtgaagcttg 1140
 ttgccacaa tggcactgag ctttcatcag accttctcaa gtcaccaaca attgaacatg 1200
 aacttgg 1207

<210> 15
 <211> 1546
 <212> DNA
 <213> *Haliotis tuberculata*

<220>

<223> HtH2-h

<400> 15

```

agccacaga ggaccagttg aagaaacaga agtcactcgc caacatactg acggcaatgc 60
acactttcat cgtaaggaag ttgattcgct gtccctggat gaagcaaca acttgaagaa 120
tgccctttac aagctacaga acgaccacag tctaaccgga tacgaagcaa tctctgggta 180
ccatggatac cccaatctgt gtccggaaga aggcgatgac aaaatacccc tgctgcgtcc 240
ccggatgggc atctttcctt actggcacag actcttgacc attcaactgg aaagagctct 300
tgagcacaat ggtgcactgc ttggtgttcc ttactgggac tggacaagg accgtgcgtc 360
actgccggcg ttcttctccg actccagcaa caacaatccc tacttcaagt accacatcgc 420
cgggtgttggc cagcagaccg tcagagagcc aactagtctt atatataacc agcccaaat 480
ccatggttat gattatctct attacctagc attgaccacg cttgaagaaa acaattactg 540
ggactttgag gttcagtatg agatcctcca caacgccgtc cactcctggc ttggaggatc 600
ccagaagtat tccatgtcta ccctggagta ttccggcctt gaccctgtct ttatgatcct 660
tcactcgggt ctagacagac tttggatcat ctggcaagaa cttcagaaga tcaggagaaa 720
gccctacaac ttcgctaaat gtgcttatca tatgatggaa gagccactgg cgccttcag 780
ctatccatct atcaaccagg acgagttcac ccgtgccaac tccaagcctt ctacagtttt 840
tgacagccat aagttcggct accattacga taacctgaat gttagagggtc acagcatcca 900
agaactcaac acaatcatca atgacttgag aaacacagac agaactctac caggatttgt 960
tttgtcaggc atcggtagct ctgctagtgt caagatctat ctccgaacag atgacaatga 1020
cgaagaagtt ggaactttca ctgtcctggg aggagagagg gaaatgccat gggcctacga 1080
gcgagttttc aagtatgaca tcacagaggt tgcagataga cttaaaatta agttatgggg 1140
cacccttta acttccggaa ctggagatca catccttacg aatggaatcg gtggtaaa 1200
gagcctacc caaatccttt catcatctac agacctgcca atcatgacta cgatgttctt 1260
gttatcccag tanggaagaa accttcacat cctcccaaa gttgtcgtca agaaaggcac 1320
ccgcacgag ttccaccag tcgatgattc agttacgaga ccagttgttg atcttgggaa 1380
ctacactgca ctcttcaact gtgtgggtacc accgttcaca taccacggat tcgaactgaa 1440
ccacgtctat tctgtcaagc ctggtgacta ctatgttact ggaccacga gagacctttg 1500
ccagaatgca gatgtcagga ttcatatcca tgttgaggat gagtaa 1546

```

<210> 16

<211> 967

<212> DNA

<213> Megathura crenulata

<220>

<223> KLH1-b

<400> 16

```

ggcctaccgt actgggactg gactgaaccc atgacacaca ttccgggtct ggcaggaaac 60
aaaaattatg tggattctca tgggtgcctc cacacaaatc cttttcatag ttcagtgtat 120
gcatttgaag aaaatgctcc ccacacaaaa agacaaatag atcaaagact ctttaaacc 180
gctacctttg gacaccacac agacctgttc aaccagattt tgtatgcctt tgaacaagaa 240
gattactgtg actttgaagt ccaatttgag attaccata acacgattca cgcttggaca 300
ggaggaagcg aacatttctc aatgtcgtcc ctacattaca cagctttcga tcctttgttt 360
tactttcacc attctaactg tgatcgtctt tgggccgttt ggcaagcctt acagatgaga 420
cggcataaac cctacagggc cactgcgcgc atatctcttg aacatatgca tctgaaacca 480
ttcgcccttt catctccctt taacaataac gaaaagactc atgccaatgc catgccaaac 540
aagatctacg actatgaaaa tgtcctccat tacacatacg aagatttaac atttggaggc 600
atctctctgg aaacataga aaagatgatc cacgaaaacc agcaagaaga cagaatatat 660
gccggttttc tctggctgg catacgtact tcagcaaatg ttgatattct cattaaaact 720
accgattccg tgcaacataa ggctggaaca tttgcagtgc tcggtggaag caaggaaatg 780
aagtggggat ttgatcgcgt tttcaagttt gacatcacgc acgttttgaa agatctcgat 840
ctcactgctg atggcgattt cgaagttact gttgacatca ctgaagtcga tggaaactaaa 900
cttgcataca gtcttattcc acatgcttct gtcattcgtg agcatgcacg tggttaagctg 960
aataagag 967

```

<210> 17

<211> 1242

<212> DNA

<213> Megathura crenulata

<220>

<223> KLH1-c

<400> 17

ttaaatttga	caaagtgcc	aggagtcgtc	ttattcgaaa	aatgtagac	cgtttgagcc	60
ccgaggagat	gaatgaactt	cgtaaagccc	tagccttact	gaaagaggac	aaaagtgccg	120
gtggatttca	gcagcttggt	gcattccatg	gggagccaaa	atgggtgcct	agtcccgaag	180
catctaaaaa	atttgcctgc	tgtgttcacg	gcatgtctgt	gttccctcac	tggcatcgac	240
tgttgacggg	tcagagtga	aatgctttga	gacgacatgg	ctacgatgga	gctttgccgt	300
actgggattg	gacctctcct	cttaatcacc	ttcccgaact	ggcagatcat	gagaagtacg	360
tcgaccctga	agatggggta	gagaagcata	acccttggtt	cgatgggtcat	atagatacag	420
tcgacaaaa	aacaacaaga	agtgttcaga	ataaactctt	cgaacagcct	gagtttggtc	480
attatacaag	cattgccaaa	caagtactgc	tagcgttgga	acaggacaat	ttctgtgact	540
ttgaaatcca	atatgagatt	gccataaact	acatccatgc	acttgttagga	ggcgctcagc	600
cttatgggtat	ggcatcgctt	cgctacactg	cttttgatcc	actattctac	ttgcatcact	660
ctaatacaga	tcgtatatgg	gcaatatggc	aggctttaca	gaagtacaga	ggaaaaccgt	720
acaacgttgc	taactgtgct	gttacatcga	tgagagaacc	tttgcaacca	tttggcctct	780
ctgccaatat	caacacagac	catgtaacca	aggagcattc	agtgccattc	aacgtttttg	840
attacaagac	caatttcaat	tatgaatatg	acactttgga	atttaacggg	ctctcaatct	900
ctcagttgaa	taaaaagctc	gaagcgataa	agagccaaga	cagggtcttt	gcaggcttcc	960
tgttatctgg	tttcaagaaa	tcattctctg	ttaaattcaa	tatttgcacc	gatagcagca	1020
actgtcacc	cgctggagag	ttttaccttc	tgggtgatga	aaacgagatg	ccatgggcat	1080
acgatagagt	cttcaaatat	gacataaccg	aaaaactcca	cgatctaaag	ctgcatgcag	1140
aagaccactt	ctacattgac	tatgaagtat	ttgaccttaa	accagcaagc	ctgggaaaag	1200
atttgttcaa	gcagccttca	gtcattcatg	aaccaagaat	ag		1242

<210> 18

<211> 1236

<212> DNA

<213> Megathura crenulata

<220>

<223> KLH1-d

<400> 18

gtcaccatga	aggcgaagta	tatcaagctg	aagtaacttc	tgccaaccgt	attcgaaaaa	60
acattgaaaa	tctgagcctt	ggtgaactcg	aaagtctgag	agctgccttc	ctggaaattg	120
aaaacgatgg	aacttacgaa	tcaatagcta	aattccatgg	tagccctggg	ttgtgccagt	180
taaatggtaa	ccccatctct	tgttgtgtcc	atggcatgcc	aactttccct	cactggcaca	240
gactgtacgt	ggttgtcgtt	gagaatgccc	tcctgaaaaa	aggatcatct	gtagctgttc	300
cctattggga	ctggacaaaa	cgaatcgaac	atttacctca	cctgatttca	gacgccactt	360
actacaattc	caggcaacat	cactatgaga	caaaccatt	ccatcatggc	aaaatcacac	420
acgagaatga	aatcactact	agggatccca	aggacagcct	cttccattca	gactactttt	480
acgagcaggt	cctttacgcc	ttggagcagg	ataacttctg	tgatttcgag	attcagttgg	540
agatattaca	caatgcattg	cattctttac	ttggtggcaa	aggtaaatat	tccatgtcaa	600
accttgatta	cgctgctttt	gatcctgtgt	tcttccttca	tcacgcaacg	actgacagaa	660
tctgggcaat	ctggcaagac	cttcagagggt	tccgaaaacg	gccataccga	gaagcgaatt	720
gcgctatcca	attgatgcac	acgccactcc	agccgtttga	taagagcgac	aacaatgacg	780
aggcaacgaa	aacgcacatg	actccacatg	atggttttga	atatcaaaa	agctttgggt	840
atgcttacga	taatctggaa	ctgaatcact	actcgattcc	tcagcttgat	cacatcttgc	900
aagaaagaaa	aaggcatgac	agagtattcg	ctggcttcc	ccttcacaat	attggaacat	960
ctgccgatgg	ccatgtattt	gtatgtctcc	caactgggga	acacacgaag	gactgcagtc	1020
atgaggctgg	tatgttctcc	atcttagggc	gtcaaacgga	gatgtccttt	gtatttgaca	1080
gactttacaa	acttgacata	actaaagcct	tgaaaaagaa	cggtgtgcac	ctgcaagggg	1140
atttcgatct	ggaaattgag	attacggctg	tgaatggatc	tcattctagac	agtcattgtca	1200
tccactctcc	cactatactg	tttgaggccg	gaacag			1236

<210> 19

<211> 241

<212> DNA

<213> Megathura crenulata

<220>

<223> KLH1-e

<400> 19

attctgcccc	cacagatgat	ggacacactg	aaccagtgat	gattcgcaaa	gatattcacac	60
aattggacaa	gcgtcaacaa	ctgtcactgg	tgaagccct	cgagtcctatg	aaagccgacc	120
attcatctga	tgggttccag	gcaatcgctt	ccttccatgc	tcttctcct	ctttgtccat	180
caccagctgc	ttcaaagagg	tttgctgtgt	gcgtccatgg	catgcccaacc	ttcccgaat	240

g

241

<210> 20
 <211> 949
 <212> DNA
 <213> Megathura crenulata

<220>
 <223> KLH2-b

<400> 20
 ggcctgccct actgggattg gaccatgcc aatgagtcatt tgccagaact ggctacaagt 60
 gagacctacc tcgatccagt tactggggaa actaaaaaca accctttcca tcacgcccac 120
 gtggcggtttg aaaatggtgt aacaagcagg aatcctgatg ccaaactttt tatgaaacca 180
 acttacggag accacactta cctcttcgac agcatgatct acgcatttga gcaggaagac 240
 ttctgcgact ttgaagtcca atatgagctc acgcataatg caatacatgc atgggttgga 300
 ggcagtga aaagtattcaat gtcttctctt cactacactg cttttgatcc tatattttac 360
 ctccatcact caaatgttga tcgtctcttg gccatttggc aagctcttca aatcaggaga 420
 ggcaagtctt acaaggccca ctgctcctcg tctcaagaaa gagaaccatt aaagcctttt 480
 gcattcagtt cccactgaa caacaacgag aaaacgtacc acaactctgt cccactaac 540
 gtttatgact atgtgggagt tttgcactat cgatatgatg accttcagtt tggcggtatg 600
 accatgtcag aacttgagga atatattcac aagcagacac aacatgatag aacctttgca 660
 gattcttcc tttcatatat tggaaacatca gcaagcgtag atatcttcat caatcgagaa 720
 gtcatgata aatacaaaagt gggaaagttt gtagtacttg gtggatccaa agaaatgaaa 780
 tggggcctttg atagaatgta caagtatgag atcactgagg ctctgaagac gctgaatgtt 840
 gcagtggatg atgggttcag cattactgtt gagatcaccc atgttgatgg atctcccca 900
 tctgcagatc tcattccacc tcctgctata atctttgaac gtggtcatg 949

<210> 21
 <211> 760
 <212> DNA
 <213> Megathura crenulata

<220>
 <223> KLH2-c

<400> 21
 ctgatgccaa agactttggc catagcagaa aaatcaggaa agccgttgat tctctgacag 60
 tcgaagaaca aacttcgttg aggcgagcta tggcagatct acaggacgac aaaacatcag 120
~~ggggtttcca geagattgca geatttcaag gagaaccaca atgggttcca agccccgaag 180~~
 cggagaaaaa atttgcattg tgtgttcatg gaatggctgt tttccctcac tggcacagat 240
 tgctgacagt tcaaggagaa aatgctctga ggaaacatgg ctttactggt ggactgccct 300
 actgggactg gactcgatca atgagcgccc ttccacattt tgttgctgat cctacttaca 360
 atgatgctat ttccagccag gaagaagata acccatggca tcatggtcac atagactctg 420
 ttgggcatga tactacaaga gatgtgcgtg atgatcttta tcaatctcct ggtttcggtc 480
 actacacaga tattgcacaa caagtccttc tggcctttga gcaggacagt ttctgtgatt 540
 ttgaggtaca atttgaaatt gcccataatt tcatacatgc actgattggt ggtaacgaac 600
 catacagtat gtcacttttg aggtatacta catacgtacc aatcttcttc ttgcaccact 660
 ccagtagaca ccgacttttg gccatctggc aagcaatcac tagtgccggc gcctgcaggt 720
 cgaccataag ggagagctcc caacgcgttg gatgcaatct 760

<210> 22
 <211> 323
 <212> DNA
 <213> Megathura crenulata

<220>
 <223> KLH2-d

<400> 22
 gttcacacca ggctgatgaa tatcgtgagg cagtaacaag cgctagccac ataagaaaaa 60
 atatccggga cctctcagag ggagaaattg agagcatcag atctgctttc ctccaaattc 120
 aaaaagaggg tatatatgaa aacattgcaa agttccatgg aaaaccagga ctttgtgaac 180
 atgatggaca tcctgttgct tgttgtgtcc atggcatgcc cacctttccc cactggcaca 240
 gactgtacgt tcttcagggtg gagaatgcgc tcttagaacg aggtctgca gttgctgttc 300

cttactggga ctggacccta cct

323

<210> 23

<211> 988

<212> DNA

<213> Megathura crenulata

<220>

<223> KLH2-g

<400> 23

```

atggctgtgt ttccgcactg gcacagactg tttgtgaaac agatggagga cgcacttgtct 60
gctcatggag ctcatattgg cataccatac tgggattgga caagtgcgtt tagtcatctg 120
cccgccctag tgactgacca cgagaacaat cccttccacc acggccatat tggatcatctg 180
aatgtggata catctcgatc tccaagagac atgctgttta atgacatctga acaaggctca 240
gaatcattct tctacagaca gggttctctg actctagaac agacagactt ctgccaattt 300
gaagttcagt ttgaacttac acacaatgcc atccactctt ggactggagg acatactcca 360
tatggaatgt catcactgga atatacagca tatgatccac tcttttatct ccaccattcc 420
aacactgacg gtatctgggc catctggcag gcactccaga aatatagagg tcttccatac 480
aacgcagctc actgcgatat ccaagttctg aaacaacctc ttaaaccatt cagcgagtcc 540
aggaatccaa acccagtcac cagagccaat tctagggccg ttgattcatt tgattatgag 600
aaattcaatt atcaatatga cacacttacc ttccacggac tttctatccc agaacttgat 660
cccatgcttc aagagagaaa gaaggaagag agaacatttg cagccttctt gttgcacgga 720
ttggcgcca gtgctgatgt ttcgtttgat gtctgcacac ctgatgggtca ttgtgccttt 780
gctggaacct tcgcggtact tgggtggggag cttgagatgc cctgggtcctt tgaaagattg 840
ttccggttac atatacaaaa gggttctcaag cagatgaatc ttcactatga ttctgagttc 900
cactttgagt tgaagattgt tggcacagat ggaacagaac tgccatcgga tcgtatcaag 960
agccctacca ttgaacacca tggaggag                                     988

```

<210> 24

<211> 310

<212> DNA

<213> Megathura crenulata

<220>

<223> KLH2-h

<400> 24

```

gtcacgatca cagtgaacgt cacgatggat ttttcaggaa ggaagtcggt tccctgtccc 60
tggatgaagc caatgacctt aaaaatgcac tgtacaagct gcagaatgat cagggtccca 120
atggatatga atcaatagcc ggttaccatg gctatccatt cctctgccc gaacatgggtg 180
aagaccagta cgcagtctgt gtccacggaa tgccctgtatt tccacattgg cacagacttc 240
atacaatcca gtttgagaga gctctcaaag aacatggttc tcatttgggt ctgccatact 300
gggactggac                                     310

```

<210> 25

<211> 422

<212> PRT

<213> Haliotis tuberculata

<220>

<221> SIGNAL

<222> (1)..(15)

<220>

<223> HtH1-a

<400> 25

```

Leu Val Gln Phe Leu Leu Val Ala Leu Val Ala Gly Ala Gly Ala Asp
 1               5               10              15

Asn Val Val Arg Lys Asp Val Ser His Leu Thr Asp Asp Glu Val Gln
      20              25              30

Ala Leu His Gly Ala Leu His Asp Val Thr Ala Ser Thr Gly Pro Leu
 35              40              45

```

Ser Phe Glu Asp Ile Thr Ser Tyr His Ala Ala Pro Ala Ser Cys Asp
 50 55 60
 Tyr Lys Gly Arg Lys Ile Ala Cys Cys Val His Gly Met Pro Ser Phe
 65 70 75 80
 Pro Phe Trp His Arg Ala Tyr Val Val Gln Ala Glu Arg Ala Leu Leu
 85 90 95
 Ser Lys Arg Lys Thr Val Gly Met Pro Tyr Trp Asp Trp Thr Gln Thr
 100 105 110
 Leu Thr His Leu Pro Ser Leu Val Thr Glu Pro Ile Tyr Ile Asp Ser
 115 120 125
 Lys Gly Gly Lys Ala Gln Thr Asn Tyr Trp Tyr Arg Gly Glu Ile Ala
 130 135 140
 Phe Ile Asn Lys Lys Thr Ala Arg Ala Val Asp Asp Arg Leu Phe Glu
 145 150 155 160
 Lys Val Glu Pro Gly His Tyr Thr His Leu Met Glu Thr Val Leu Asp
 165 170 175
 Ala Leu Glu Gln Asp Glu Phe Cys Lys Phe Glu Ile Gln Phe Glu Leu
 180 185 190
 Ala His Asn Ala Ile His Tyr Leu Val Gly Gly Lys Phe Glu Tyr Ser
 195 200 205
 Met Ser Asn Leu Glu Tyr Thr Ser Tyr Asp Pro Ile Phe Phe Leu His
 210 215 220
 His Ser Asn Val Asp Arg Leu Phe Ala Ile Trp Gln Arg Leu Gln Glu
 225 230 235 240
 Leu Arg Gly Lys Asn Pro Asn Ala Met Asp Cys Ala His Glu Leu Ala
 245 250 255

~~His Gln Gln Leu Gln Pro Phe Asn Arg Asp Ser Asn Pro Val Gln Leu~~
~~260 265 270~~
 Thr Lys Asp His Ser Thr Pro Ala Asp Leu Phe Asp Tyr Lys Gln Leu
 275 280 285
 Gly Tyr Ser Tyr Asp Ser Leu Asn Leu Asn Gly Met Thr Pro Glu Gln
 290 295 300
 Leu Lys Thr Glu Leu Asp Glu Arg His Ser Lys Glu Arg Ala Phe Ala
 305 310 315 320
 Ser Phe Arg Leu Ser Gly Phe Gly Gly Ser Ala Asn Val Val Val Tyr
 325 330 335
 Ala Cys Val Pro Asp Asp Asp Pro Arg Ser Asp Asp Tyr Cys Glu Lys
 340 345 350
 Ala Gly Asp Phe Phe Ile Leu Gly Gly Gln Ser Glu Met Pro Trp Arg
 355 360 365
 Phe Tyr Arg Pro Phe Phe Tyr Asp Val Thr Glu Ala Val His His Leu
 370 375 380
 Gly Val Pro Leu Ser Gly His Tyr Tyr Val Lys Thr Glu Leu Phe Ser
 385 390 395 400

Val Asn Gly Thr Ala Leu Ser Pro Asp Leu Leu Pro Gln Pro Thr Val
 405 410 415

Ala Tyr Arg Pro Gly Lys
 420

<210> 26
 <211> 419
 <212> PRT
 <213> *Haliotis tuberculata*

<220>
 <223> HtH1-b

<400> 26
 Gly His Leu Asp Pro Pro Val His His Arg His Asp Asp Asp Leu Ile
 1 5 10 15

Val Arg Lys Asn Ile Asp His Leu Thr Arg Glu Glu Glu Tyr Glu Leu
 20 25 30

Arg Met Ala Leu Glu Arg Phe Gln Ala Asp Thr Ser Val Asp Gly Tyr
 35 40 45

Gln Ala Thr Val Glu Tyr His Gly Leu Pro Ala Arg Cys Pro Arg Pro
 50 55 60

Asp Ala Lys Val Arg Phe Ala Cys Cys Met His Gly Met Ala Ser Phe
 65 70 75 80

Pro His Trp His Arg Leu Phe Val Thr Gln Val Glu Asp Ala Leu Val
 85 90 95

Arg Arg Gly Ser Pro Ile Gly Val Pro Tyr Trp Asp Trp Thr Lys Pro
 100 105 110

Met Thr His Leu Pro Asp Leu Ala Ser Asn Glu Thr Tyr Val Asp Pro
 115 120 125

Tyr Gly His Thr His His Asn Pro Phe Phe Asn Ala Asn Ile Ser Phe
 130 135 140

Glu Glu Gly His His His Thr Ser Arg Met Ile Asp Ser Lys Leu Phe
 145 150 155 160

Ala Pro Val Ala Phe Gly Glu His Ser His Leu Phe Asp Gly Ile Leu
 165 170 175

Tyr Ala Phe Glu Gln Glu Asp Phe Cys Asp Phe Glu Ile Gln Phe Glu
 180 185 190

Leu Val His Asn Ser Ile His Ala Trp Ile Gly Gly Ser Glu Asp Tyr
 195 200 205

Ser Met Ala Thr Leu His Tyr Thr Ala Phe Asp Pro Ile Phe Tyr Leu
 210 215 220

His His Ser Asn Val Asp Arg Leu Trp Ala Ile Trp Gln Ala Leu Gln
 225 230 235 240

Ile Arg Arg His Lys Pro Tyr Gln Ala His Cys Ala Gln Ser Val Glu
 245 250 255

Gln Leu Pro Met Lys Pro Phe Ala Phe Pro Ser Pro Leu Asn Asn Asn

```
<210> 27
<211> 414
<212> PRT
<213> Haliotis tuberculata

<220>
<223> HtH1-c
```

<400>	27																
Pro	Thr	Phe	Glu	Asp	Glu	Lys	His	Ser	Leu	Arg	Ile	Arg	Lys	Asn	Val		
1				5					10					15			
Asp	Ser	Leu	Thr	Pro	Glu	Glu	Thr	Asn	Glu	Leu	Arg	Lys	Ala	Leu	Glu		
			20					25					30				
Leu	Leu	Glu	Asn	Asp	His	Thr	Ala	Gly	Gly	Phe	Asn	Gln	Leu	Gly	Ala		
		35					40					45					
Phe	His	Gly	Glu	Pro	Lys	Trp	Cys	Pro	Asn	Pro	Glu	Ala	Glu	His	Lys		
	50					55					60						
Val	Ala	Cys	Cys	Val	His	Gly	Met	Ala	Val	Phe	Pro	His	Trp	His	Arg		
65					70					75					80		
Leu	Leu	Ala	Leu	Gln	Ala	Glu	Asn	Ala	Leu	Arg	Lys	His	Gly	Tyr	Ser		
				85					90					95			
Gly	Ala	Leu	Pro	Tyr	Trp	Asp	Trp	Thr	Arg	Pro	Leu	Ser	Gln	Leu	Pro		
			100					105					110				
Asp	Leu	Val	Ser	His	Glu	Gln	Tyr	Thr	Asp	Pro	Ser	Asp	His	His	Val		
		115					120					125					

Lys His Asn Pro Trp Phe Asn Gly His Ile Asp Thr Val Asn Gln Asp
 130 135 140
 Thr Thr Arg Ser Val Arg Glu Asp Leu Tyr Gln Gln Pro Glu Phe Gly
 145 150 155 160
 His Phe Thr Asp Ile Ala Gln Gln Val Leu Leu Ala Leu Glu Gln Asp
 165 170 175
 Asp Phe Cys Ser Phe Glu Val Gln Tyr Glu Ile Ser His Asn Phe Ile
 180 185 190
 His Ala Leu Val Gly Gly Thr Asp Ala Tyr Gly Met Ala Ser Leu Arg
 195 200 205
 Tyr Thr Ala Tyr Asp Pro Ile Phe Phe Leu His His Ser Asn Thr Asp
 210 215 220
 Arg Ile Trp Ala Ile Trp Gln Ser Leu Gln Lys Tyr Arg Gly Lys Pro
 225 230 235 240
 Tyr Asn Thr Ala Asn Cys Ala Ile Glu Ser Met Arg Arg Pro Leu Gln
 245 250 255
 Pro Phe Gly Leu Ser Ser Ala Ile Asn Pro Asp Arg Ile Thr Arg Glu
 260 265 270
 His Ala Ile Pro Phe Asp Val Phe Asn Tyr Arg Asp Asn Leu His Tyr
 275 280 285
 Val Tyr Asp Thr Leu Glu Phe Asn Gly Leu Ser Ile Ser Gln Leu Asp
 290 295 300
 Arg Glu Leu Glu Lys Ile Lys Ser His Glu Arg Val Phe Ala Gly Phe
 305 310 315 320
 Leu Leu Ser Gly Ile Lys Lys Ser Ala Leu Val Lys Phe Glu Val Cys
 325 330 335
 Thr Pro Pro Asp Asn Cys His Lys Ala Gly Glu Phe Tyr Leu Leu Gly
 340 345 350
 Asp Glu Asn Glu Met Ala Trp Ala Tyr Asp Arg Leu Phe Lys Tyr Asp
 355 360 365
 Ile Thr Gln Val Leu Glu Ala Asn His Leu His Phe Tyr Asp His Leu
 370 375 380
 Phe Ile Arg Tyr Glu Val Phe Asp Leu Lys Gly Val Ser Leu Gly Thr
 385 390 395 400
 Asp Leu Phe His Thr Ala Asn Val Val His Asp Ser Gly Thr
 405 410

<210> 28
 <211> 413
 <212> PRT
 <213> Haliotis tuberculata

<220>
 <223> HtH1-d

<400> 28
 Gly Thr Arg Asp Arg Asp Asn Tyr Val Glu Glu Val Thr Gly Ala Ser
 1 5 10 15

His Ile Arg Lys Asn Leu Asn Asp Leu Asn Thr Gly Glu Met Glu Ser
 20 25 30
 Leu Arg Ala Ala Phe Leu His Ile Gln Asp Asp Gly Thr Tyr Glu Ser
 35 40 45
 Ile Ala Gln Tyr His Gly Lys Pro Gly Lys Cys Gln Leu Asn Asp His
 50 55 60
 Asn Ile Ala Cys Cys Val His Gly Met Pro Thr Phe Pro Gln Trp His
 65 70 75 80
 Arg Leu Tyr Val Val Gln Val Glu Asn Ala Leu Leu Asn Arg Gly Ser
 85 90 95
 Gly Val Ala Val Pro Tyr Trp Glu Trp Thr Ala Pro Ile Asp His Leu
 100 105 110
 Pro His Phe Ile Asp Asp Ala Thr Tyr Phe Asn Ser Arg Gln Gln Arg
 115 120 125
 Tyr Asp Pro Asn Pro Phe Phe Arg Gly Lys Val Thr Phe Glu Asn Ala
 130 135 140
 Val Thr Thr Arg Asp Pro Gln Ala Gly Leu Phe Asn Ser Asp Tyr Met
 145 150 155 160
 Tyr Glu Asn Val Leu Leu Ala Leu Glu Gln Glu Asn Tyr Cys Asp Phe
 165 170 175
 Glu Ile Gln Phe Glu Leu Val His Asn Ala Leu His Ser Met Leu Gly
 180 185 190
 Gly Lys Gly Gln Tyr Ser Met Ser Ser Leu Asp Tyr Ser Ala Phe Asp
 195 200 205
 Pro Val Phe Phe Leu His His Ala Asn Thr Asp Arg Leu Trp Ala Ile
 210 215 220

 Trp Gln Glu Leu Gln Arg Phe Arg Glu Leu Pro Tyr Glu Glu Ala Asn
 225 230 235 240
 Cys Ala Ile Asn Leu Met His Gln Pro Leu Lys Pro Phe Ser Asp Pro
 245 250 255
 His Glu Asn His Asp Asn Val Thr Leu Lys Tyr Ser Lys Pro Gln Asp
 260 265 270
 Gly Phe Asp Tyr Gln Asn His Phe Gly Tyr Lys Tyr Asp Asn Leu Glu
 275 280 285
 Phe His His Leu Ser Ile Pro Ser Leu Asp Ala Thr Leu Lys Gln Arg
 290 295 300
 Arg Asn His Asp Arg Val Phe Ala Gly Phe Leu Leu His Asn Ile Gly
 305 310 315 320
 Thr Ser Ala Asp Ile Thr Ile Tyr Ile Cys Leu Pro Asp Gly Arg Arg
 325 330 335
 Gly Asn Asp Cys Ser His Glu Ala Gly Thr Phe Tyr Ile Leu Gly Gly
 340 345 350
 Glu Thr Glu Met Pro Phe Ile Phe Asp Arg Leu Tyr Lys Phe Glu Ile
 355 360 365

Thr Lys Pro Leu Gln Gln Leu Gly Val Lys Leu His Gly Gly Val Phe
370 375 380

Glu Leu Glu Leu Glu Ile Lys Ala Tyr Asn Gly Ser Tyr Leu Asp Pro
385 390 395 400

His Thr Phe Asp Pro Thr Ile Ile Phe Glu Pro Gly Thr
405 410

<210> 29

<211> 420

<212> PRT

<213> *Haliotis tuberculata*

<220>

<223> HtH1-e

<400> 29

Asp Thr His Ile Leu Asp His Asp His Glu Glu Glu Ile Leu Val Arg
1 5 10 15

Lys Asn Ile Ile Asp Leu Ser Pro Arg Glu Arg Val Ser Leu Val Lys
20 25 30

Ala Leu Gln Arg Met Lys Asn Asp Arg Ser Ala Asp Gly Tyr Gln Ala
35 40 45

Ile Ala Ser Phe His Ala Leu Pro Pro Leu Cys Pro Asn Pro Ser Ala
50 55 60

Ala His Arg Tyr Ala Cys Cys Val His Gly Met Ala Thr Phe Pro Gln
65 70 75 80

Trp His Arg Leu Tyr Thr Val Gln Val Gln Asp Ala Leu Arg Arg His
85 90 95

Gly Ser Leu Val Gly Ile Pro Tyr Trp Asp Trp Thr Lys Pro Val Asn
100 105 110

Glu Leu Pro Glu Leu Leu Ser Ser Ala Thr Phe Tyr His Pro Ile Arg
115 120 125

Asn Ile Asn Ile Ser Asn Pro Phe Leu Gly Ala Asp Ile Glu Phe Glu
130 135 140

Gly Pro Gly Val His Thr Glu Arg His Ile Asn Thr Glu Arg Leu Phe
145 150 155 160

His Ser Gly Asp His Asp Gly Tyr His Asn Trp Phe Phe Glu Thr Val
165 170 175

Leu Phe Ala Leu Glu Gln Glu Asp Tyr Cys Asp Phe Glu Ile Gln Phe
180 185 190

Glu Ile Ala His Asn Gly Ile His Thr Trp Ile Gly Gly Ser Ala Val
195 200 205

Tyr Gly Met Gly His Leu His Tyr Ala Ser Tyr Asp Pro Ile Phe Tyr
210 215 220

Ile His His Ser Gln Thr Asp Arg Ile Trp Ala Ile Trp Gln Glu Leu
225 230 235 240

Gln Lys Tyr Arg Gly Leu Ser Gly Ser Glu Ala Asn Cys Ala Ile Glu

245 250 255
 His Met Arg Thr Pro Leu Lys Pro Phe Ser Phe Gly Pro Pro Tyr Asn
 260 265 270
 Leu Asn Ser His Thr Gln Glu Tyr Ser Lys Pro Glu Asp Thr Phe Asp
 275 280 285
 Tyr Lys Lys Phe Gly Tyr Arg Tyr Asp Ser Leu Glu Leu Glu Gly Arg
 290 295 300
 Ser Ile Ser Arg Ile Asp Glu Leu Ile Gln Gln Arg Gln Glu Lys Asp
 305 310 315 320
 Arg Thr Phe Ala Gly Phe Leu Leu Lys Gly Phe Gly Thr Ser Ala Ser
 325 330 335
 Val Ser Leu Gln Val Cys Arg Val Asp His Thr Cys Lys Asp Ala Gly
 340 345 350
 Tyr Phe Thr Ile Leu Gly Gly Ser Ala Glu Met Pro Trp Ala Phe Asp
 355 360 365
 Arg Leu Tyr Lys Tyr Asp Ile Thr Lys Thr Leu His Asp Met Asn Leu
 370 375 380
 Arg His Glu Asp Thr Phe Ser Ile Asp Val Thr Ile Thr Ser Tyr Asn
 385 390 395 400
 Gly Thr Val Leu Ser Gly Asp Leu Ile Gln Thr Pro Ser Ile Ile Phe
 405 410 415
 Val Pro Gly Arg
 420

<210> 30
 <211> 417
 <212> PRT
 <213> *Haliotis tuberculata*

<220>
 <223> HtH1-f

<400> 30
 His Lys Leu Asn Ser Arg Lys His Thr Pro Asn Arg Val Arg His Glu
 1 5 10 15
 Leu Ser Ser Leu Ser Ser Arg Asp Ile Ala Ser Leu Lys Ala Ala Leu
 20 25 30
 Thr Ser Leu Gln His Asp Asn Gly Thr Asp Gly Tyr Gln Ala Ile Ala
 35 40 45
 Ala Phe His Gly Val Pro Ala Gln Cys His Glu Pro Ser Gly Arg Glu
 50 55 60
 Ile Ala Cys Cys Ile His Gly Met Ala Thr Phe Pro His Trp His Arg
 65 70 75 80
 Leu Tyr Thr Leu Gln Leu Glu Gln Ala Leu Arg Arg His Gly Ser Ser
 85 90 95
 Val Ala Val Pro Tyr Trp Asp Trp Thr Lys Pro Ile Thr Glu Leu Pro
 100 105 110

His Ile Leu Thr Asp Gly Glu Tyr Tyr Asp Val Trp Gln Asn Ala Val
 115 120 125
 Leu Ala Asn Pro Phe Ala Arg Gly Tyr Val Lys Ile Lys Asp Ala Phe
 130 135 140
 Thr Val Arg Asn Val Gln Glu Ser Leu Phe Lys Met Ser Ser Phe Gly
 145 150 155 160
 Lys His Ser Leu Leu Phe Asp Gln Ala Leu Leu Ala Leu Glu Gln Thr
 165 170 175
 Asp Tyr Cys Asp Phe Glu Val Gln Phe Glu Val Met His Asn Thr Ile
 180 185 190
 His Tyr Leu Val Gly Gly Arg Gln Thr Tyr Ala Phe Ser Ser Leu Glu
 195 200 205
 Tyr Ser Ser Tyr Asp Pro Ile Phe Phe Ile His His Ser Phe Val Asp
 210 215 220
 Lys Ile Trp Ala Val Trp Gln Glu Leu Gln Ser Arg Arg His Leu Gln
 225 230 235 240
 Phe Arg Thr Ala Asp Cys Ala Val Gly Leu Met Gly Gln Ala Met Arg
 245 250 255
 Pro Phe Asn Lys Asp Phe Asn His Asn Ser Phe Thr Lys Lys His Ala
 260 265 270
 Val Pro Asn Thr Val Phe Asp Tyr Glu Asp Leu Gly Tyr Asn Tyr Asp
 275 280 285
 Asn Leu Glu Ile Ser Gly Leu Asn Leu Asn Glu Ile Glu Ala Leu Ile
 290 295 300
 Ala Lys Arg Lys Ser His Ala Arg Val Phe Ala Gly Phe Leu Leu Phe
 305 310 315 320
 Gly Leu Gly Thr Ser Ala Asp Ile His Leu Glu Ile Cys Lys Thr Ser
 325 330 335

 Glu Asn Cys His Asp Ala Gly Val Ile Phe Ile Leu Gly Gly Ser Ala
 340 345 350
 Glu Met His Trp Ala Tyr Asn Arg Leu Tyr Lys Tyr Asp Ile Thr Glu
 355 360 365
 Ala Leu Gln Glu Phe Asp Ile Asn Pro Glu Asp Val Phe His Ala Asp
 370 375 380
 Glu Pro Phe Phe Leu Arg Leu Ser Val Val Ala Val Asn Gly Thr Val
 385 390 395 400
 Ile Pro Ser Ser His Leu His Gln Pro Thr Ile Ile Tyr Glu Pro Gly
 405 410 415
 Glu

<210> 31
 <211> 403
 <212> PRT
 <213> Haliotis tuberculata

<220>

<223> HtH1-g

<400> 31

Asp His His Asp Asp His Gln Ser Gly Ser Ile Ala Gly Ser Gly Val
 1 5 10 15

Arg Lys Asp Val Asn Thr Leu Thr Lys Ala Glu Thr Asp Asn Leu Arg
 20 25 30

Glu Ala Leu Trp Gly Val Met Ala Asp His Gly Pro Asn Gly Phe Gln
 35 40 45

Ala Ile Ala Ala Phe His Gly Lys Pro Ala Leu Cys Pro Met Pro Asp
 50 55 60

Gly His Asn Tyr Ser Cys Cys Thr His Gly Met Ala Thr Phe Pro His
 65 70 75 80

Trp His Arg Leu Tyr Thr Lys Gln Met Glu Asp Ala Met Arg Ala His
 85 90 95

Gly Ser His Val Gly Leu Pro Tyr Trp Asp Trp Thr Ala Ala Phe Thr
 100 105 110

His Leu Pro Thr Leu Val Thr Asp Thr Asp Asn Asn Pro Phe Gln His
 115 120 125

Gly His Ile Asp Tyr Leu Asn Val Ser Thr Thr Arg Ser Pro Arg Asp
 130 135 140

Met Leu Phe Asn Asp Pro Glu His Gly Ser Glu Ser Phe Phe Tyr Arg
 145 150 155 160

Gln Val Leu Leu Ala Leu Glu Gln Thr Asp Phe Cys Lys Phe Glu Val
 165 170 175

Gln Phe Glu Ile Thr His Asn Ala Ile His Ser Trp Thr Gly Gly His
 180 185 190

Ser Pro Tyr Gly Met Ser Thr Leu Asp Phe Thr Ala Tyr Asp Pro Leu
 195 200 205

Phe Trp Leu His His Ser Asn Thr Asp Arg Ile Trp Ala Val Trp Gln
 210 215 220

Ala Leu Gln Glu Tyr Arg Gly Leu Pro Tyr Asn His Ala Asn Cys Glu
 225 230 235 240

Ile Gln Ala Met Lys Thr Pro Leu Arg Pro Phe Ser Asp Asp Ile Asn
 245 250 255

His Asn Pro Val Thr Lys Ala Asn Ala Lys Pro Leu Asp Val Phe Glu
 260 265 270

Tyr Asn Arg Leu Ser Phe Gln Tyr Asp Asn Leu Ile Phe His Gly Tyr
 275 280 285

Ser Ile Pro Glu Leu Asp Arg Val Leu Glu Glu Arg Lys Glu Glu Asp
 290 295 300

Arg Ile Phe Ala Ala Phe Leu Leu Ser Gly Ile Lys Arg Ser Ala Asp
 305 310 315 320

Val Val Phe Asp Ile Cys Gln Pro Glu His Glu Cys Val Phe Ala Gly
 325 330 335

Thr Phe Ala Ile Leu Gly Gly Glu Leu Glu Met Pro Trp Ser Phe Asp
 340 345 350

Arg Leu Phe Arg Tyr Asp Ile Thr Lys Val Met Lys Gln Leu His Leu
 355 360 365

Arg His Asp Ser Asp Phe Thr Phe Arg Val Lys Ile Val Gly Thr Asp
 370 375 380

Asp His Glu Leu Pro Ser Asp Ser Val Lys Ala Pro Thr Ile Glu Phe
 385 390 395 400

Glu Pro Gly

<210> 32
 <211> 511
 <212> PRT
 <213> Haliotis tuberculata

<220>
 <223> HtH1-h

<400> 32
 Val His Arg Gly Gly Asn His Glu Asp Glu His His Asp Asp Arg Leu
 1 5 10 15

Ala Asp Val Leu Ile Arg Lys Glu Val Asp Phe Leu Ser Leu Gln Glu
 20 25 30

Ala Asn Ala Ile Lys Asp Ala Leu Tyr Lys Leu Gln Asn Asp Asp Ser
 35 40 45

Lys Gly Gly Phe Glu Ala Ile Ala Gly Tyr His Gly Tyr Pro Asn Met
 50 55 60

Cys Pro Glu Arg Gly Thr Asp Lys Tyr Pro Cys Cys Val His Gly Met
 65 70 75 80

Pro Val Phe Pro His Trp His Arg Leu His Thr Ile Gln Met Glu Arg
 85 90 95

Ala Leu Lys Asn His Gly Ser Pro Met Gly Ile Pro Tyr Trp Asp Trp
 100 105 110

Thr Lys Lys Met Ser Ser Leu Pro Ser Phe Phe Gly Asp Ser Ser Asn
 115 120 125

Asn Asn Pro Phe Tyr Lys Tyr Tyr Ile Arg Gly Val Gln His Glu Thr
 130 135 140

Thr Arg Asp Val Asn Gln Arg Leu Phe Asn Gln Thr Lys Phe Gly Glu
 145 150 155 160

Phe Asp Tyr Leu Tyr Tyr Leu Thr Leu Gln Val Leu Glu Glu Asn Ser
 165 170 175

Tyr Cys Asp Phe Glu Val Gln Tyr Glu Ile Leu His Asn Ala Val His
 180 185 190

Ser Trp Leu Gly Gly Thr Gly Gln Tyr Ser Met Ser Thr Leu Glu Tyr
 195 200 205

Ser Ala Phe Asp Pro Val Phe Met Ile His His Ser Ser Leu Asp Arg

210	215	220
Ile Trp Ile Leu Trp Gln Lys Leu Gln Lys Ile Arg Met Lys Pro Tyr 225 230 235 240		
Tyr Ala Leu Asp Cys Ala Gly Asp Arg Leu Met Lys Asp Pro Leu His 245 250 255		
Pro Phe Asn Tyr Glu Thr Val Asn Glu Asp Glu Phe Thr Arg Ile Asn 260 265 270		
Ser Phe Pro Ser Ile Leu Phe Asp His Tyr Arg Phe Asn Tyr Glu Tyr 275 280 285		
Asp Asn Met Arg Ile Arg Gly Gln Asp Ile His Glu Leu Glu Glu Val 290 295 300		
Ile Gln Glu Leu Arg Asn Lys Asp Arg Ile Phe Ala Gly Phe Val Leu 305 310 315 320		
Ser Gly Leu Arg Ile Ser Ala Thr Val Lys Val Phe Ile His Ser Lys 325 330 335		
Asn Asp Thr Ser His Glu Glu Tyr Ala Gly Glu Phe Ala Val Leu Gly 340 345 350		
Gly Glu Lys Glu Met Pro Trp Ala Tyr Glu Arg Met Leu Lys Leu Asp 355 360 365		
Ile Ser Asp Ala Val His Lys Leu His Val Lys Asp Glu Asp Ile Arg 370 375 380		
Phe Arg Val Val Val Thr Ala Tyr Asn Gly Asp Val Val Thr Thr Arg 385 390 395 400		
Leu Ser Gln Pro Phe Ile Val His Arg Pro Ala His Val Ala His Asp 405 410 415		
Ile Leu Val Ile Pro Val Gly Ala Gly His Asp Leu Pro Pro Lys Val 420 425 430		
Val Val Lys Ser Gly Thr Lys Val Glu Phe Thr Pro Ile Asp Ser Ser 435 440 445		
Val Asn Lys Ala Met Val Glu Leu Gly Ser Tyr Thr Ala Met Ala Lys 450 455 460		
Cys Ile Val Pro Pro Phe Ser Tyr His Gly Phe Glu Leu Asp Lys Val 465 470 475 480		
Tyr Ser Val Asp His Gly Asp Tyr Tyr Ile Ala Ala Gly Thr His Ala 485 490 495		
Leu Cys Glu Gln Asn Leu Arg Leu His Ile His Val Glu His Glu 500 505 510		

<210> 33
 <211> 334
 <212> PRT
 <213> Haliotis tuberculata

<220>
 <223> HtH2-b

<400> 33

His Arg Leu Phe Val Thr Gln Val Glu Asp Ala Leu Ile Arg Arg Gly
 1 5 10 15
 Ser Pro Ile Gly Val Pro Tyr Trp Asp Trp Thr Gln Pro Met Ala His
 20 25 30
 Leu Pro Gly Leu Ala Asp Asn Ala Thr Tyr Arg Asp Pro Ile Ser Gly
 35 40 45
 Asp Ser Arg His Asn Pro Phe His Asp Val Glu Val Ala Phe Glu Asn
 50 55 60
 Gly Arg Thr Glu Arg His Pro Asp Ser Arg Leu Phe Glu Gln Pro Leu
 65 70 75 80
 Phe Gly Lys His Thr Arg Leu Phe Asp Ser Ile Val Tyr Ala Phe Glu
 85 90 95
 Gln Glu Asp Phe Cys Asp Phe Glu Val Gln Phe Glu Met Thr His Asn
 100 105 110
 Asn Ile His Ala Trp Ile Gly Gly Gly Glu Lys Tyr Ser Met Ser Ser
 115 120 125
 Leu His Tyr Thr Ala Phe Asp Pro Ile Phe Tyr Leu Arg His Ser Asn
 130 135 140
 Thr Asp Arg Leu Trp Ala Ile Trp Gln Ala Leu Gln Ile Arg Arg Asn
 145 150 155 160
 Arg Pro Tyr Lys Ala His Cys Ala Trp Ser Glu Glu Arg Gln Pro Leu
 165 170 175
 Lys Pro Phe Ala Phe Ser Ser Pro Leu Asn Asn Asn Glu Lys Thr Tyr
 180 185 190
 Glu Asn Ser Val Pro Thr Asn Val Tyr Asp Tyr Glu Gly Val Leu Gly
 195 200 205
 Tyr Thr Tyr Asp Asp Leu Asn Phe Gly Gly Met Asp Leu Gly Gln Leu
 210 215 220

Glu Glu Tyr Ile Gln Arg Gln Arg Gln Arg Asp Arg Thr Phe Ala Gly
 225 230 235 240
 Phe Phe Leu Ser His Ile Gly Thr Ser Ala Asn Val Glu Ile Ile Ile
 245 250 255
 Asp His Gly Thr Leu His Thr Ser Val Gly Thr Phe Ala Val Leu Gly
 260 265 270
 Gly Glu Lys Glu Met Lys Trp Gly Phe Asp Arg Leu Tyr Lys Tyr Glu
 275 280 285
 Ile Thr Asp Glu Leu Arg Gln Leu Asn Leu Arg Ala Asp Asp Val Phe
 290 295 300
 Ser Ile Ser Val Lys Val Thr Asp Val Asp Gly Ser Glu Leu Ser Ser
 305 310 315 320
 Glu Leu Ile Pro Ser Ala Ala Ile Ile Phe Glu Arg Ser His
 325 330

<212> PRT

<213> Haliotis tuberculata

<220>

<223> HtH2-c

<400> 34

Ile Asp His Gln Asp Pro His His Asp Thr Ile Ile Arg Lys Asn Val
 1 5 10 15

Asp Asn Leu Thr Pro Glu Glu Ile Asn Ser Leu Arg Arg Ala Met Ala
 20 25 30

Asp Leu Gln Ser Asp Lys Thr Ala Gly Gly Phe Gln Gln Ile Ala Ala
 35 40 45

Phe His Gly Glu Pro Lys Trp Cys Pro Ser Pro Asp Ala Glu Lys Lys
 50 55 60

Phe Ser Cys Cys Val His Gly Met Ala Val Phe Pro His Trp His Arg
 65 70 75 80

Leu Leu Thr Val Gln Gly Glu Asn Ala Leu Arg Lys His Gly Cys Leu
 85 90 95

Gly Ala Leu Pro Tyr Trp Asp Trp Thr Arg Pro Leu Ser His Leu Pro
 100 105 110

Asp Leu Val Leu Val Ser Ser Arg Thr Thr Pro Met Pro Tyr Ser Thr
 115 120 125

Val Glu Ala Arg Asn Pro Trp Tyr Ser Gly His Ile Asp Thr Val Gly
 130 135 140

Val Asp Thr Thr Arg Ser Val Arg Gln Glu Leu Tyr Glu Ala Pro Gly
 145 150 155 160

Phe Gly His Tyr Thr Gly Val Ala Lys Gln Val Leu Leu Ala Leu Glu
 165 170 175

Gln Asp Asp Phe Cys Asp Phe Glu Val Gln Phe Glu Ile Ala His Asn
 180 185 190

Phe Ile His Ala Leu Val Gly Gly Ser Glu Pro Tyr Gly Met Ala Ser
 195 200 205

Leu Arg Tyr Thr Thr Tyr Asp Pro Ile Phe Tyr Leu His His Ser Asn
 210 215 220

Thr Asp Arg Leu Trp Ala Ile Trp Gln Ala Leu Gln Lys Tyr Arg Gly
 225 230 235 240

Lys Pro Tyr Asn Ser Ala Asn Cys Ala Ile Ala Ser Met Arg Lys Pro
 245 250 255

Leu Gln Pro Phe Gly Leu Thr Asp Glu Ile Asn Pro Asp Asp Glu Thr
 260 265 270

Arg Gln His Ala Val Pro Phe Ser Val Phe Asp Tyr Lys Asn Asn Phe
 275 280 285

Asn Tyr Glu Tyr Asp Thr Leu Asp Phe Asn Gly Leu Ser Ile Ser Gln
 290 295 300

Leu Asp Arg Glu Leu Ser Arg Arg Lys Ser His Asp Arg Val Phe Ala
 305 310 315 320


```
<210> 35
<211> 415
<212> PRT
<213> Haliotis tuberculata
```

<220>
<223> HtH2-d

<400> 35
Gly His His Gln Ala Asp Glu Tyr Asp Glu Val Val Thr Ala Ala Ser
1 5 10 15
His Ile Arg Lys Asn Leu Lys Asp Leu Ser Lys Gly Glu Val Glu Ser
20 25 30
Leu Arg Ser Ala Phe Leu Gln Leu Gln Asn Asp Gly Val Tyr Glu Asn
35 40 45

Ile Ala Lys Phe His Gly Lys Pro Gly Leu Cys Asp Asp Asn Gly Arg
50 55 60

Lys Val Ala Cys Cys Val His Gly Met Pro Thr Phe Pro Gln Trp His
65 70 75 80

Arg Leu Tyr Val Leu Gln Val Glu Asn Ala Leu Leu Glu Arg Gly Ser
85 90 95

Ala Val Ser Val Pro Tyr Trp Asp Trp Thr Glu Thr Phe Thr Glu Leu
100 105 110

Pro Ser Leu Ile Ala Glu Ala Thr Tyr Phe Asn Ser Arg Gln Gln Thr
115 120 125

Phe Asp Pro Asn Pro Phe Phe Arg Gly Lys Ile Ser Phe Glu Asn Ala
130 135 140

Val	Thr	Thr	Arg	Asp	Pro	Gln	Pro	Glu	Leu	Tyr	Val	Asn	Arg	Tyr	Tyr
145					150					155					160

Tyr Gln Asn Val Met Leu Val Phe Glu Gln Asp Asn Tyr Cys Asp Phe
165 170 175

Glu Ile Gln Phe Glu Met Val His Asn Val Leu His Ala Trp Leu Gly

180	185	190
Gly Arg Ala Thr Tyr Ser Ile 195	Ser Ser Leu Asp Tyr 200	Ser Ala Phe Asp 205
Pro Val Phe Phe Leu His 210	His Ala Asn Thr Asp 215	Arg Leu Trp Ala Ile 220
Trp Gln Glu Leu Gln Arg 225	Tyr Arg Lys Lys Pro 230	Tyr Asn Glu Ala Asp 235 240
Cys Ala Ile Asn Leu Met 245	Arg Lys Pro Leu His 250	Pro Phe Asp Asn Ser 255
Asp Leu Asn His Asp 260	Pro Val Thr Phe Lys 265	Tyr Ser Lys Pro Thr Asp 270
Gly Phe Asp Tyr Gln Asn 275	Asn Phe Gly Tyr Lys 280	Tyr Asp Asn Leu Glu 285
Phe Asn His Phe Ser Ile 290	Pro Arg Leu Glu Glu 295	Ile Ile Arg Ile Arg 300
Gln Arg Gln Asp Arg Val 305	Phe Ala Gly Phe Leu 310	Leu His Asn Ile Gly 315 320
Thr Ser Ala Thr Val 325	Glu Ile Phe Val Cys 330	Val Pro Thr Thr Ser Gly 335
Glu Gln Asn Cys Glu Asn 340	Lys Ala Gly Thr Phe 345	Ala Val Leu Gly Gly 350
Glu Thr Glu Met Ala Phe 355	His Phe Asp Arg Leu 360	Tyr Arg Phe Asp Ile 365
Ser Glu Thr Leu Arg Asp 370	Leu Gly Ile Gln Leu 375	Asp Ser His Asp Phe 380
Asp Leu Ser Ile Lys Ile 385	Gln Gly Val Asn Gly 390	Ser Tyr Leu Asp Pro 395 400

His Ile Leu Pro Glu Pro 405	Ser Leu Ile Phe Val 410	Pro Gly Ser Ser 415
--------------------------------	----------------------------	------------------------

<210> 36

<211> 418

<212> PRT

<213> Haliotis tuberculata

<220>

<223> HtH2-e

<400> 36

Ser Phe Leu Arg Pro Asp Gly His Ser Asp Asp Ile Leu Val Arg Lys
1 5 10 15

Glu Val Asn Ser Leu Thr Thr Arg Glu Thr Ala Ser Leu Ile His Ala
20 25 30

Leu Lys Ser Met Gln Glu Asp His Ser Pro Asp Gly Phe Gln Ala Ile
35 40 45

Ala Ser Phe His Ala Leu Pro Pro Leu Cys Pro Ser Pro Ser Ala Ala
50 55 60

His Arg Tyr Ala Cys Cys Val His Gly Met Ala Thr Phe Pro Gln Trp
 65 70 75 80
 His Arg Leu Tyr Thr Val Gln Phe Gln Asp Ala Leu Arg Arg His Gly
 85 90 95
 Ala Thr Val Gly Val Pro Tyr Trp Asp Trp Leu Arg Pro Gln Ser His
 100 105 110
 Leu Pro Glu Leu Val Thr Met Glu Thr Tyr His Asp Ile Trp Ser Asn
 115 120 125
 Arg Asp Phe Pro Asn Pro Phe Tyr Gln Ala Asn Ile Glu Phe Glu Gly
 130 135 140
 Glu Asn Ile Thr Thr Glu Arg Glu Val Ile Ala Asp Lys Leu Phe Val
 145 150 155 160
 Lys Gly Gly His Val Phe Asp Lys Leu Val Leu Gln Thr Ser His Pro
 165 170 175
 Ser Ala Glu Gln Glu Asn Tyr Cys Asp Phe Glu Ile Gln Phe Glu Ile
 180 185 190
 Leu His Asn Gly Val His Thr Trp Val Gly Gly Ser Arg Thr Tyr Ser
 195 200 205
 Ile Gly His Leu His Tyr Ala Phe Tyr Asp Pro Leu Phe Tyr Leu His
 210 215 220
 His Phe Gln Thr Asp Arg Ile Trp Ala Ile Trp Gln Glu Leu Gln Glu
 225 230 235 240
 Gln Arg Gly Leu Ser Gly Asp Glu Ala His Cys Ala Leu Glu Gln Met
 245 250 255
 Arg Glu Pro Leu Lys Pro Phe Ser Phe Gly Ala Pro Tyr Asn Trp Asn
 260 265 270
 Gln Leu Thr Gln Asp Phe Ser Arg Pro Glu Asp Thr Phe Asp Tyr Arg
 275 280 285

Lys Phe Gly Tyr Glu Tyr Asp Asn Leu Glu Phe Leu Gly Met Ser Val
 290 295 300
 Ala Glu Leu Asp Gln Tyr Ile Ile Glu His Gln Glu Asn Asp Arg Val
 305 310 315 320
 Phe Ala Gly Phe Leu Leu Ser Gly Phe Gly Gly Ser Ala Ser Val Asn
 325 330 335
 Phe Gln Val Cys Arg Ala Asp Ser Thr Cys Gln Asp Ala Gly Tyr Phe
 340 345 350
 Thr Val Leu Gly Gly Ser Ala Glu Met Ala Trp Ala Phe Asp Arg Leu
 355 360 365
 Tyr Lys Tyr Asp Ile Thr Glu Thr Leu Glu Lys Met His Leu Arg Tyr
 370 375 380
 Asp Asp Asp Phe Thr Ile Ser Val Ser Leu Thr Ala Asn Asn Gly Thr
 385 390 395 400
 Val Leu Ser Ser Ser Leu Ile Pro Thr Pro Ser Val Ile Phe Gln Arg
 405 410 415

Gly His

<210> 37

<211> 416

<212> PRT

<213> *Haliotis tuberculata*

<220>

<223> HtH2-f

<400> 37

Arg Asp Ile Asn Thr Arg Ser Met Ser Pro Asn Arg Val Arg Arg Glu
 1 5 10 15

Leu Ser Asp Leu Ser Ala Arg Asp Leu Ser Ser Leu Lys Ser Ala Leu
 20 25 30

Arg Asp Leu Gln Glu Asp Asp Gly Pro Asn Gly Tyr Gln Ala Leu Ala
 35 40 45

Ala Phe His Gly Leu Pro Ala Gly Cys His Asp Ser Arg Gly Asn Glu
 50 55 60

Ile Ala Cys Cys Ile His Gly Met Pro Thr Phe Pro Gln Trp His Arg
 65 70 75 80

Leu Tyr Thr Leu Gln Leu Glu Met Ala Leu Arg Arg His Gly Ser Ser
 85 90 95

Val Ala Ile Pro Tyr Trp Asp Trp Thr Lys Pro Ile Ser Glu Leu Pro
 100 105 110

Ser Leu Phe Thr Ser Pro Glu Tyr Tyr Asp Pro Trp His Asp Ala Val
 115 120 125

Val Asn Asn Pro Phe Ser Lys Gly Phe Val Lys Phe Ala Asn Thr Tyr
 130 135 140

Thr Val Arg Asp Pro Gln Glu Met Leu Phe Gln Leu Cys Glu His Gly
 145 150 155 160

Glu Ser Ile Leu Tyr Glu Gln Thr Leu Leu Ala Leu Glu Gln Thr Asp
 165 170 175

Tyr Cys Asp Phe Glu Val Gln Phe Glu Val Leu His Asn Val Ile His
 180 185 190

Tyr Leu Val Gly Gly Arg Gln Thr Tyr Ala Leu Ser Ser Leu His Tyr
 195 200 205

Ala Ser Tyr Asp Pro Phe Phe Phe Ile His His Ser Phe Val Asp Lys
 210 215 220

Met Trp Val Val Trp Gln Ala Leu Gln Lys Arg Arg Lys Leu Pro Tyr
 225 230 235 240

Lys Arg Ala Asp Cys Ala Val Asn Leu Met Thr Lys Pro Met Arg Pro
 245 250 255

Phe Asp Ser Asp Met Asn Gln Asn Pro Phe Thr Lys Met His Ala Val
 260 265 270

Pro Asn Thr Leu Tyr Asp Tyr Glu Thr Leu Tyr Tyr Ser Tyr Asp Asn
 275 280 285

Leu Glu Ile Gly Gly Arg Asn Leu Asp Gln Leu Gln Ala Glu Ile Asp
290 295 300

Arg Ser Arg Ser His Asp Arg Val Phe Ala Gly Phe Leu Leu Arg Gly
305 310 315 320

Ile Gly Thr Ser Ala Asp Val Arg Phe Trp Ile Cys Arg Asn Glu Asn
325 330 335

Asp Cys His Arg Gly Gly Ile Ile Phe Ile Leu Gly Gly Ala Lys Glu
340 345 350

Met Pro Trp Ser Phe Asp Arg Asn Phe Lys Phe Asp Ile Thr His Val
355 360 365

Leu Glu Asn Ala Gly Ile Ser Pro Glu Asp Val Phe Asp Ala Glu Glu
370 375 380

Pro Phe Tyr Ile Lys Val Glu Ile His Ala Val Asn Lys Thr Met Ile
385 390 395 400

Pro Ser Ser Val Ile Pro Ala Pro Thr Ile Ile Tyr Ser Pro Gly Glu
405 410 415

<210> 38

<211> 402

<212> PRT

<213> Haliotis tuberculata

<220>

<223> HtH2-g

<400> 38

Gly Arg Ala Ala Asp Ser Ala His Ser Ala Asn Ile Ala Gly Ser Gly
1 5 10 15

Val Arg Lys Asp Val Thr Thr Leu Thr Val Ser Glu Thr Glu Asn Leu
20 25 30

Arg Gln Ala Leu Gln Gly Val Ile Asp Asp Thr Gly Pro Asn Gly Tyr
35 40 45

Gln Ala Ile Ala Ser Phe His Gly Ser Pro Pro Met Cys Glu Met Asn
50 55 60

Gly Arg Lys Val Ala Cys Cys Ala His Gly Met Ala Ser Phe Pro His
65 70 75 80

Trp His Arg Leu Tyr Val Lys Gln Met Glu Asp Ala Leu Ala Asp His
85 90 95

Gly Ser His Ile Gly Ile Pro Tyr Trp Asp Trp Thr Thr Ala Phe Thr
100 105 110

Glu Leu Pro Ala Leu Val Thr Asp Ser Glu Asn Asn Pro Phe His Glu
115 120 125

Gly Arg Ile Asp His Leu Gly Val Thr Thr Ser Arg Ser Pro Arg Asp
130 135 140

Met Leu Phe Asn Asp Pro Glu Gln Gly Ser Glu Ser Phe Phe Tyr Arg

145 150 155 160
 Gln Val Leu Leu Ala Leu Glu Gln Thr Asp Tyr Cys Gln Phe Glu Val
 165 170 175
 Gln Phe Glu Leu Thr His Asn Ala Ile His Ser Trp Thr Gly Gly Arg
 180 185 190
 Ser Pro Tyr Gly Met Ser Thr Leu Glu Phe Thr Ala Tyr Asp Pro Leu
 195 200 205
 Phe Trp Leu His His Ser Asn Thr Asp Arg Ile Trp Ala Val Trp Gln
 210 215 220
 Ala Leu Gln Lys Tyr Arg Gly Leu Pro Tyr Asn Glu Ala His Cys Glu
 225 230 235 240
 Ile Gln Val Leu Lys Gln Pro Leu Arg Pro Phe Asn Asp Asp Ile Asn
 245 250 255
 His Asn Pro Ile Thr Lys Thr Asn Ala Arg Pro Ile Asp Ser Phe Asp
 260 265 270
 Tyr Glu Arg Phe Asn Tyr Gln Tyr Asp Thr Leu Ser Phe His Gly Lys
 275 280 285
 Ser Ile Pro Glu Leu Asn Asp Leu Leu Glu Glu Arg Lys Arg Glu Glu
 290 295 300
 Arg Thr Phe Ala Ala Phe Leu Leu Arg Gly Ile Gly Cys Ser Ala Asp
 305 310 315 320
 Val Val Phe Asp Ile Cys Arg Pro Asn Gly Asp Cys Val Phe Ala Gly
 325 330 335
 Thr Phe Ala Val Leu Gly Gly Glu Leu Glu Met Pro Trp Ser Phe Asp
 340 345 350
 Arg Leu Phe Arg Tyr Asp Ile Thr Arg Val Met Asn Gln Leu His Leu
 355 360 365

Gln Tyr Asp Ser Asp Phe Ser Phe Arg Val Lys Leu Val Ala Thr Asn
 370 375 380
 Gly Thr Glu Leu Ser Ser Asp Leu Leu Lys Ser Pro Thr Ile Glu His
 385 390 395 400

Glu Leu

<210> 39
 <211> 515
 <212> PRT
 <213> Haliotis tuberculata

<220>
 <223> HtH2-h

<400> 39
 Gly Ala His Arg Gly Pro Val Glu Glu Thr Glu Val Thr Arg Gln His
 1 5 10 15

Thr Asp Gly Asn Ala His Phe His Arg Lys Glu Val Asp Ser Leu Ser
 20 25 30

Leu Asp Glu Ala Asn Asn Leu Lys Asn Ala Leu Tyr Lys Leu Gln Asn
35 40 45

Asp His Ser Leu Thr Gly Tyr Glu Ala Ile Ser Gly Tyr His Gly Tyr
50 55 60

Pro Asn Leu Cys Pro Glu Glu Gly Asp Asp Lys Ile Pro Leu Leu Arg
65 70 75 80

Pro Arg Met Gly Ile Phe Pro Tyr Trp His Arg Leu Leu Thr Ile Gln
85 90 95

Leu Glu Arg Ala Leu Glu His Asn Gly Ala Leu Leu Gly Val Pro Tyr
100 105 110

Trp Asp Trp Asn Lys Asp Leu Ser Ser Leu Pro Ala Phe Phe Ser Asp
115 120 125

Ser Ser Asn Asn Asn Pro Tyr Phe Lys Tyr His Ile Ala Gly Val Gly
130 135 140

His Asp Thr Val Arg Glu Pro Thr Ser Leu Ile Tyr Asn Gln Pro Gln
145 150 155 160

Ile His Gly Tyr Asp Tyr Leu Tyr Tyr Leu Ala Leu Thr Thr Leu Glu
165 170 175

Glu Asn Asn Tyr Trp Asp Phe Glu Val Gln Tyr Glu Ile Leu His Asn
180 185 190

Ala Val His Ser Trp Leu Gly Gly Ser Gln Lys Tyr Ser Met Ser Thr
195 200 205

Leu Glu Tyr Ser Ala Phe Asp Pro Val Phe Met Ile Leu His Ser Gly
210 215 220

Leu Asp Arg Leu Trp Ile Ile Trp Gln Glu Leu Gln Lys Ile Arg Arg
225 230 235 240

~~Lys Pro Tyr Asn Phe Ala Lys Cys Ala Tyr His Met Met Glu Glu Pro~~
~~245 250 255~~

Leu Ala Pro Phe Ser Tyr Pro Ser Ile Asn Gln Asp Glu Phe Thr Arg
260 265 270

Ala Asn Ser Lys Pro Ser Thr Val Phe Asp Ser His Lys Phe Gly Tyr
275 280 285

His Tyr Asp Asn Leu Asn Val Arg Gly His Ser Ile Gln Glu Leu Asn
290 295 300

Thr Ile Ile Asn Asp Leu Arg Asn Thr Asp Arg Ile Tyr Ala Gly Phe
305 310 315 320

Val Leu Ser Gly Ile Gly Thr Ser Ala Ser Val Lys Ile Tyr Leu Arg
325 330 335

Thr Asp Asp Asn Asp Glu Glu Val Gly Thr Phe Thr Val Leu Gly Gly
340 345 350

Glu Arg Glu Met Pro Trp Ala Tyr Glu Arg Val Phe Lys Tyr Asp Ile
355 360 365

Thr Glu Val Ala Asp Arg Leu Lys Ile Lys Leu Trp Gly His Pro Leu
370 375 380

Thr Ser Gly Thr Gly Asp His Ile Leu Thr Asn Gly Ile Gly Gly Lys
385 390 395 400

Gln Glu Pro Thr Gln Ile Leu Ser Ser Ser Thr Asp Leu Pro Ile Met
405 410 415

Thr Thr Met Phe Leu Leu Ser Gln Xaa Gly Arg Asn Leu His Ile Pro
420 425 430

Pro Lys Val Val Val Lys Lys Gly Thr Arg Ile Glu Phe His Pro Val
435 440 445

Asp Asp Ser Val Thr Arg Pro Val Val Asp Leu Gly Ser Tyr Thr Ala
450 455 460

Leu Phe Asn Cys Val Val Pro Pro Phe Thr Tyr His Gly Phe Glu Leu
465 470 475 480

Asn His Val Tyr Ser Val Lys Pro Gly Asp Tyr Tyr Val Thr Gly Pro
485 490 495

Thr Arg Asp Leu Cys Gln Asn Ala Asp Val Arg Ile His Ile His Val
500 505 510

Glu Asp Glu
515

<210> 40

<211> 322

<212> PRT

<213> Megathura crenulata

<220>

<223> KLH1-b

<400> 40

Gly Leu Pro Tyr Trp Asp Trp Thr Glu Pro Met Thr His Ile Pro Gly
1 5 10 15

Leu Ala Gly Asn Lys Thr Tyr Val Asp Ser His Gly Ala Ser His Thr
20 25 30

Asn Pro Phe His Ser Ser Val Ile Ala Phe Glu Glu Asn Ala Pro His
35 40 45

Thr Lys Arg Gln Ile Asp Gln Arg Leu Phe Lys Pro Ala Thr Phe Gly
50 55 60

His His Thr Asp Leu Phe Asn Gln Ile Leu Tyr Ala Phe Glu Gln Glu
65 70 75 80

Asp Tyr Cys Asp Phe Glu Val Gln Phe Glu Ile Thr His Asn Thr Ile
85 90 95

His Ala Trp Thr Gly Gly Ser Glu His Phe Ser Met Ser Ser Leu His
100 105 110

Tyr Thr Ala Phe Asp Pro Leu Phe Tyr Phe His His Ser Asn Val Asp
115 120 125

Arg Leu Trp Ala Val Trp Gln Ala Leu Gln Met Arg Arg His Lys Pro
130 135 140

Tyr Arg Ala His Cys Ala Ile Ser Leu Glu His Met His Leu Lys Pro
145 150 155 160

Phe Ala Phe Ser Ser Pro Leu Asn Asn Asn Glu Lys Thr His Ala Asn
165 170 175

Ala Met Pro Asn Lys Ile Tyr Asp Tyr Glu Asn Val Leu His Tyr Thr
180 185 190

Tyr Glu Asp Leu Thr Phe Gly Gly Ile Ser Leu Glu Asn Ile Glu Lys
195 200 205

Met Ile His Glu Asn Gln Gln Glu Asp Arg Ile Tyr Ala Gly Phe Leu
210 215 220

Leu Ala Gly Ile Arg Thr Ser Ala Asn Val Asp Ile Phe Ile Lys Thr
225 230 235 240

Thr Asp Ser Val Gln His Lys Ala Gly Thr Phe Ala Val Leu Gly Gly
245 250 255

Ser Lys Glu Met Lys Trp Gly Phe Asp Arg Val Phe Lys Phe Asp Ile
260 265 270

Thr His Val Leu Lys Asp Leu Asp Leu Thr Ala Asp Gly Asp Phe Glu
275 280 285

Val Thr Val Asp Ile Thr Glu Val Asp Gly Thr Lys Leu Ala Ser Ser
290 295 300

Leu Ile Pro His Ala Ser Val Ile Arg Glu His Ala Arg Gly Lys Leu
305 310 315 320

Asn Arg

<210> 41

<211> 414

<212> PRT

<213> *Megathura crenulata*

<220>

<223> KLH1-c

<400> 41

Val Lys Phe Asp Lys Val Pro Arg Ser Arg Leu Ile Arg Lys Asn Val
1 5 10 15

Asp Arg Leu Ser Pro Glu Glu Met Asn Glu Leu Arg Lys Ala Leu Ala
20 25 30

Leu Leu Lys Glu Asp Lys Ser Ala Gly Gly Phe Gln Gln Leu Gly Ala
35 40 45

Phe His Gly Glu Pro Lys Trp Cys Pro Ser Pro Glu Ala Ser Lys Lys
50 55 60

Phe Ala Cys Cys Val His Gly Met Ser Val Phe Pro His Trp His Arg
65 70 75 80

Leu Leu Thr Val Gln Ser Glu Asn Ala Leu Arg Arg His Gly Tyr Asp
85 90 95

Gly Ala Leu Pro Tyr Trp Asp Trp Thr Ser Pro Leu Asn His Leu Pro
100 105 110

Glu Leu Ala Asp His Glu Lys Tyr Val Asp Pro Glu Asp Gly Val Glu
 115 120 125
 Lys His Asn Pro Trp Phe Asp Gly His Ile Asp Thr Val Asp Lys Thr
 130 135 140
 Thr Thr Arg Ser Val Gln Asn Lys Leu Phe Glu Gln Pro Glu Phe Gly
 145 150 155 160
 His Tyr Thr Ser Ile Ala Lys Gln Val Leu Leu Ala Leu Glu Gln Asp
 165 170 175
 Asn Phe Cys Asp Phe Glu Ile Gln Tyr Glu Ile Ala His Asn Tyr Ile
 180 185 190
 His Ala Leu Val Gly Gly Ala Gln Pro Tyr Gly Met Ala Ser Leu Arg
 195 200 205
 Tyr Thr Ala Phe Asp Pro Leu Phe Tyr Leu His His Ser Asn Thr Asp
 210 215 220
 Arg Ile Trp Ala Ile Trp Gln Ala Leu Gln Lys Tyr Arg Gly Lys Pro
 225 230 235 240
 Tyr Asn Val Ala Asn Cys Ala Val Thr Ser Met Arg Glu Pro Leu Gln
 245 250 255
 Pro Phe Gly Leu Ser Ala Asn Ile Asn Thr Asp His Val Thr Lys Glu
 260 265 270
 His Ser Val Pro Phe Asn Val Phe Asp Tyr Lys Thr Asn Phe Asn Tyr
 275 280 285
 Glu Tyr Asp Thr Leu Glu Phe Asn Gly Leu Ser Ile Ser Gln Leu Asn
 290 295 300
 Lys Lys Leu Glu Ala Ile Lys Ser Gln Asp Arg Phe Phe Ala Gly Phe
 305 310 315 320
 Leu Leu Ser Gly Phe Lys Lys Ser Ser Leu Val Lys Phe Asn Ile Cys
 325 330 335
 Thr Asp Ser Ser Asn Cys His Pro Ala Gly Glu Phe Tyr Leu Leu Gly
 340 345 350
 Asp Glu Asn Glu Met Pro Trp Ala Tyr Asp Arg Val Phe Lys Tyr Asp
 355 360 365
 Ile Thr Glu Lys Leu His Asp Leu Lys Leu His Ala Glu Asp His Phe
 370 375 380
 Tyr Ile Asp Tyr Glu Val Phe Asp Leu Lys Pro Ala Ser Leu Gly Lys
 385 390 395 400
 Asp Leu Phe Lys Gln Pro Ser Val Ile His Glu Pro Arg Ile
 405 410

<210> 42

<211> 411

<212> PRT

<213> Megathura crenulata

<220>

<223> KLH1-d

<400> 42

Gly His His Glu Gly Glu Val Tyr Gln Ala Glu Val Thr Ser Ala Asn
1 5 10 15

Arg Ile Arg Lys Asn Ile Glu Asn Leu Ser Leu Gly Glu Leu Glu Ser
20 25 30

Leu Arg Ala Ala Phe Leu Glu Ile Glu Asn Asp Gly Thr Tyr Glu Ser
35 40 45

Ile Ala Lys Phe His Gly Ser Pro Gly Leu Cys Gln Leu Asn Gly Asn
50 55 60

Pro Ile Ser Cys Cys Val His Gly Met Pro Thr Phe Pro His Trp His
65 70 75 80

Arg Leu Tyr Val Val Val Val Glu Asn Ala Leu Leu Lys Lys Gly Ser
85 90 95

Ser Val Ala Val Pro Tyr Trp Asp Trp Thr Lys Arg Ile Glu His Leu
100 105 110

Pro His Leu Ile Ser Asp Ala Thr Tyr Tyr Asn Ser Arg Gln His His
115 120 125

Tyr Glu Thr Asn Pro Phe His His Gly Lys Ile Thr His Glu Asn Glu
130 135 140

Ile Thr Thr Arg Asp Pro Lys Asp Ser Leu Phe His Ser Asp Tyr Phe
145 150 155 160

Tyr Glu Gln Val Leu Tyr Ala Leu Glu Gln Asp Asn Phe Cys Asp Phe
165 170 175

Glu Ile Gln Leu Glu Ile Leu His Asn Ala Leu His Ser Leu Leu Gly
180 185 190

Gly Lys Gly Lys Tyr Ser Met Ser Asn Leu Asp Tyr Ala Ala Phe Asp
195 200 205

Pro Val Phe Phe Leu His His Ala Thr Thr Asp Arg Ile Trp Ala Ile
210 215 220

Trp Gln Asp Leu Gln Arg Phe Arg Lys Arg Pro Tyr Arg Glu Ala Asn
225 230 235 240

Cys Ala Ile Gln Leu Met His Thr Pro Leu Gln Pro Phe Asp Lys Ser
245 250 255

Asp Asn Asn Asp Glu Ala Thr Lys Thr His Ala Thr Pro His Asp Gly
260 265 270

Phe Glu Tyr Gln Asn Ser Phe Gly Tyr Ala Tyr Asp Asn Leu Glu Leu
275 280 285

Asn His Tyr Ser Ile Pro Gln Leu Asp His Met Leu Gln Glu Arg Lys
290 295 300

Arg His Asp Arg Val Phe Ala Gly Phe Leu Leu His Asn Ile Gly Thr
305 310 315 320

Ser Ala Asp Gly His Val Phe Val Cys Leu Pro Thr Gly Glu His Thr
325 330 335

Lys Asp Cys Ser His Glu Ala Gly Met Phe Ser Ile Leu Gly Gly Gln
340 345 350

Thr Glu Met Ser Phe Val Phe Asp Arg Leu Tyr Lys Leu Asp Ile Thr
355 360 365

Lys Ala Leu Lys Lys Asn Gly Val His Leu Gln Gly Asp Phe Asp Leu
370 375 380

Glu Ile Glu Ile Thr Ala Val Asn Gly Ser His Leu Asp Ser His Val
385 390 395 400

Ile His Ser Pro Thr Ile Leu Phe Glu Ala Gly
405 410

<210> 43

<211> 111

<212> PRT

<213> Megathura crenulata

<220>

<223> KLH1-e

<400> 43

Asp Ser Ala His Thr Asp Asp Gly His Thr Glu Pro Val Met Ile Arg
1 5 10 15

Lys Asp Ile Thr Gln Leu Asp Lys Arg Gln Gln Leu Ser Leu Val Lys
20 25 30

Ala Leu Glu Ser Met Lys Ala Asp His Ser Ser Asp Gly Phe Gln Ala
35 40 45

Ile Ala Ser Phe His Ala Leu Pro Pro Leu Cys Pro Ser Pro Ala Ala
50 55 60

Ser Lys Arg Phe Ala Cys Cys Val His Gly Met Pro Thr Phe Pro Gln
65 70 75 80

Trp His Arg Leu Tyr Thr Val Gln Phe Gln Asp Ser Leu Arg Lys His
85 90 95

Gly Ala Val Val Gly Leu Pro Tyr Trp Asp Trp Thr Leu Pro Arg
100 105 110

<210> 44

<211> 317

<212> PRT

<213> Megathura crenulata

<220>

<223> KLH2-b

<400> 44

Gly Leu Pro Tyr Trp Asp Trp Thr Met Pro Met Ser His Leu Pro Glu
1 5 10 15

Leu Ala Thr Ser Glu Thr Tyr Leu Asp Pro Val Thr Gly Glu Thr Lys
20 25 30

Asn Asn Pro Phe His His Ala Gln Val Ala Phe Glu Asn Gly Val Thr
35 40 45

Ser Arg Asn Pro Asp Ala Lys Leu Phe Met Lys Pro Thr Tyr Gly Asp
50 55 60

His Thr Tyr Leu Phe Asp Ser Met Ile Tyr Ala Phe Glu Gln Glu Asp
65 70 75 80

Phe Cys Asp Phe Glu Val Gln Tyr Glu Leu Thr His Asn Ala Ile His
85 90 95

Ala Trp Val Gly Gly Ser Glu Lys Tyr Ser Met Ser Ser Leu His Tyr
100 105 110

Thr Ala Phe Asp Pro Ile Phe Tyr Leu His His Ser Asn Val Asp Arg
115 120 125

Leu Trp Ala Ile Trp Gln Ala Leu Gln Ile Arg Arg Gly Lys Ser Tyr
130 135 140

Lys Ala His Cys Ala Ser Ser Gln Glu Arg Glu Pro Leu Lys Pro Phe
145 150 155 160

Ala Phe Ser Ser Pro Leu Asn Asn Asn Glu Lys Thr Tyr His Asn Ser
165 170 175

Val Pro Thr Asn Val Tyr Asp Tyr Val Gly Val Leu His Tyr Arg Tyr
180 185 190

Asp Asp Leu Gln Phe Gly Gly Met Thr Met Ser Glu Leu Glu Glu Tyr
195 200 205

Ile His Lys Gln Thr Gln His Asp Arg Thr Phe Ala Gly Phe Phe Leu
210 215 220

Ser Tyr Ile Gly Thr Ser Ala Ser Val Asp Ile Phe Ile Asn Arg Glu
225 230 235 240

Gly His Asp Lys Tyr Lys Val Gly Ser Phe Val Val Leu Gly Gly Ser
245 250 255

Lys Glu Met Lys Trp Gly Phe Asp Arg Met Tyr Lys Tyr Glu Ile Thr
260 265 270

Glu Ala Leu Lys Thr Leu Asn Val Ala Val Asp Asp Gly Phe Ser Ile
275 280 285

Thr Val Glu Ile Thr Asp Val Asp Gly Ser Pro Pro Ser Ala Asp Leu
290 295 300

Ile Pro Pro Pro Ala Ile Ile Phe Glu Arg Gly His Ala
305 310 315

<210> 45

<211> 411

<212> PRT

<213> Megathura crenulata

<220>

<223> KLH2-c

<400> 45

Asp Ala Lys Asp Phe Gly His Ser Arg Lys Ile Arg Lys Ala Val Asp
1 5 10 15

Ser Leu Thr Val Glu Glu Gln Thr Ser Leu Arg Arg Ala Met Ala Asp
20 25 30

Leu Gln Asp Asp Lys Thr Ser Gly Gly Phe Gln Gln Ile Ala Ala Phe
35 40 45

His Gly Glu Pro Lys Trp Cys Pro Ser Pro Glu Ala Glu Lys Lys Phe
 50 55 60
 Ala Cys Cys Val His Gly Met Ala Val Phe Pro His Trp His Arg Leu
 65 70 75 80
 Leu Thr Val Gln Gly Glu Asn Ala Leu Arg Lys His Gly Phe Thr Gly
 85 90 95
 Gly Leu Pro Tyr Trp Asp Trp Thr Arg Ser Met Ser Ala Leu Pro His
 100 105 110
 Phe Val Ala Asp Pro Thr Tyr Asn Asp Ala Ile Ser Ser Gln Glu Glu
 115 120 125
 Asp Asn Pro Trp His His Gly His Ile Asp Ser Val Gly His Asp Thr
 130 135 140
 Thr Arg Asp Val Arg Asp Asp Leu Tyr Gln Ser Pro Gly Phe Gly His
 145 150 155 160
 Tyr Thr Asp Ile Ala Gln Gln Val Leu Leu Ala Phe Glu Gln Asp Ser
 165 170 175
 Phe Cys Asp Phe Glu Val Gln Phe Glu Ile Ala His Asn Phe Ile His
 180 185 190
 Ala Leu Ile Gly Gly Asn Glu Pro Tyr Ser Met Ser Ser Leu Arg Tyr
 195 200 205
 Thr Thr Tyr Asp Pro Ile Phe Phe Leu His His Ser Ser Thr Asp Arg
 210 215 220
 Leu Trp Ala Ile Trp Gln Ala Leu Gln Lys Tyr Arg Gly Lys Pro Tyr
 225 230 235 240
 Asn Thr Ala Asn Cys Ala Ile Ala Ser Met Arg Lys Pro Leu Gln Pro
 245 250 255

Phe Gly Leu Asp Ser Val Ile Asn Pro Asp Asp Glu Thr Arg Glu His
 260 265 270
 Ser Val Pro Phe Arg Val Phe Asp Tyr Lys Asn Asn Phe Asp Tyr Glu
 275 280 285
 Tyr Glu Ser Leu Ala Phe Asn Gly Leu Ser Ile Ala Gln Leu Asp Arg
 290 295 300
 Glu Leu Gln Arg Arg Lys Ser His Asp Arg Val Phe Ala Gly Phe Leu
 305 310 315 320
 Leu His Glu Ile Gly Gln Ser Ala Lys His Asn Val Ser Asp Cys Asp
 325 330 335
 His Tyr Ala Gly Glu Phe Tyr Ile Leu Gly Asp Glu Ala Glu Met Pro
 340 345 350
 Trp Arg Tyr Asp Arg Val Tyr Lys Tyr Glu Ile Thr Gln Gln Leu His
 355 360 365
 Asp Leu Asp Leu His Val Gly Asp Asn Phe Phe Leu Lys Tyr Glu Ala
 370 375 380
 Phe Asp Leu Asn Gly Gly Ser Leu Gly Gly Ser Ile Phe Ser Gln Pro
 385 390 395 400

Ser Val Ile Phe Glu Pro Ala Ala Gly Met Phe
405 410

<210> 46
<211> 109
<212> PRT
<213> Megathura crenulata

<220>
<223> KLH2-d

<400> 46
Gly Ser His Gln Ala Asp Glu Tyr Arg Glu Ala Val Thr Ser Ala Ser
1 5 10 15
His Ile Arg Lys Asn Ile Arg Asp Leu Ser Glu Gly Glu Ile Glu Ser
20 25 30
Ile Arg Ser Ala Phe Leu Gln Ile Gln Lys Glu Gly Ile Tyr Glu Asn
35 40 45
Ile Ala Lys Phe His Gly Lys Pro Gly Leu Cys Glu His Asp Gly His
50 55 60
Pro Val Ala Cys Cys Val His Gly Met Pro Thr Phe Pro His Trp His
65 70 75 80
Arg Leu Tyr Val Leu Gln Val Glu Asn Ala Leu Leu Glu Arg Gly Ser
85 90 95
Ala Val Ala Val Pro Tyr Trp Asp Trp Thr Leu Pro Arg
100 105

<210> 47
<211> 329
<212> PRT
<213> Megathura crenulata

<220>
<223> KLH2-g

<400> 47
Met Ala Val Phe Pro His Trp His Arg Leu Phe Val Lys Gln Met Glu
1 5 10 15
Asp Ala Leu Ala Ala His Gly Ala His Ile Gly Ile Pro Tyr Trp Asp
20 25 30
Trp Thr Ser Ala Phe Ser His Leu Pro Ala Leu Val Thr Asp His Glu
35 40 45
Asn Asn Pro Phe His His Gly His Ile Gly His Leu Asn Val Asp Thr
50 55 60
Ser Arg Ser Pro Arg Asp Met Leu Phe Asn Asp Pro Glu Gln Gly Ser
65 70 75 80
Glu Ser Phe Phe Tyr Arg Gln Val Leu Leu Thr Leu Glu Gln Thr Asp
85 90 95
Phe Cys Gln Phe Glu Val Gln Phe Glu Leu Thr His Asn Ala Ile His
100 105 110

Ser Trp Thr Gly His Thr Pro Tyr Gly Met Ser Ser Leu Glu Tyr
 115 120 125
 Thr Ala Tyr Asp Pro Leu Phe Tyr Leu His His Ser Asn Thr Asp Arg
 130 135 140
 Ile Trp Ala Ile Trp Gln Ala Leu Gln Lys Tyr Arg Gly Leu Pro Tyr
 145 150 155 160
 Asn Ala Ala His Cys Asp Ile Gln Val Leu Lys Gln Pro Leu Lys Pro
 165 170 175
 Phe Ser Glu Ser Arg Asn Pro Asn Pro Val Thr Arg Ala Asn Ser Arg
 180 185 190
 Ala Val Asp Ser Phe Asp Tyr Glu Lys Phe Asn Tyr Gln Tyr Asp Thr
 195 200 205
 Leu Thr Phe His Gly Leu Ser Ile Pro Glu Leu Asp Ala Met Leu Gln
 210 215 220
 Glu Arg Lys Lys Glu Glu Arg Thr Phe Ala Ala Phe Leu Leu His Gly
 225 230 235 240
 Phe Gly Ala Ser Ala Asp Val Ser Phe Asp Val Cys Thr Pro Asp Gly
 245 250 255
 His Cys Ala Phe Ala Gly Thr Phe Ala Val Leu Gly Gly Glu Leu Glu
 260 265 270
 Met Pro Trp Ser Phe Glu Arg Leu Phe Arg Tyr Asp Ile Thr Lys Val
 275 280 285
 Leu Lys Gln Met Asn Leu His Tyr Asp Ser Glu Phe His Phe Glu Leu
 290 295 300
 Lys Ile Val Gly Thr Asp Gly Thr Glu Leu Pro Ser Asp Arg Ile Lys
 305 310 315 320
 Ser Pro Thr Ile Glu His His Gly Gly
 325

<210> 48
 <211> 103
 <212> PRT
 <213> Megathura crenulata

<220>
 <223> KLH2-h

<400> 48

Gly His Asp His Ser Glu Arg His Asp Gly Phe Phe Arg Lys Glu Val
 1 5 10 15
 Gly Ser Leu Ser Leu Asp Glu Ala Asn Asp Leu Lys Asn Ala Leu Tyr
 20 25 30
 Lys Leu Gln Asn Asp Gln Gly Pro Asn Gly Tyr Glu Ser Ile Ala Gly
 35 40 45
 Tyr His Gly Tyr Pro Phe Leu Cys Pro Glu His Gly Glu Asp Gln Tyr
 50 55 60
 Ala Cys Cys Val His Gly Met Pro Val Phe Pro His Trp His Arg Leu
 65 70 75 80


His Thr Ile Gln Phe Glu Arg Ala Leu Lys Glu His Gly Ser His Leu
85 90 95

Gly Leu Pro Tyr Trp Asp Trp
100


Patentansprüche

1. Nukleinsäure-Molekül, umfassend eine für ein Hämocyanin, eine Hämocyanin-Domäne oder ein funktionelles Fragment davon mit den immunologischen Eigenschaften wenigstens einer Domäne eines Hämocyanins kodierende Nukleinsäuresequenz, wobei die Nukleinsäuresequenz ausgewählt ist aus

(a) der Gruppe der nachfolgend angegebenen DNA-Sequenzen bzw. der ihnen entsprechenden RNA-Sequenzen:



in SEQ ID NO:1 (HtH1 Domäne a + Signalpeptid),
 SEQ ID NO:2 (HtH1 Domäne b),
 SEQ ID NO:3 (HtH1 Domäne c),
 SEQ ID NO:4 (HtH1 Domäne d),
 SEQ ID NO:5 (HtH1 Domäne e),
 SEQ ID NO:6 (HtH1 Domäne f),
 SEQ ID NO:7 (HtH1 Domäne g),
 SEQ ID NO: 8 (HtH1 Domäne h),
 SEQ ID NO:9 (partielle HtH2 Domäne b),
 SEQ ID NO:10 (HtH2 Domäne c).



SEQ ID NO:11 (HtH2 Domäne d),
 SEQ ID NO:12 (HtH2 Domäne e),
 SEQ ID NO:13 (HtH2 Domäne f),
 SEQ ID NO:14 (HtH2 Domäne g),
 SEQ ID NO:15 (HtH2 Domäne h),
 SEQ ID NO:16 (partielle KLH1 Domäne b),
 SEQ ID NO:17 (KLH1 Domäne c),
 SEQ ID NO:18 (KLH1 Domäne d),
 SEQ ID NO:19 (partielle KLH1 Domäne e),
 SEQ ID NO:20 (KLH2 Domäne b),
 SEQ ID NO:21 (KLH2 Domäne c),
 SEQ ID NO:22 (partielle KLH2 Domäne d),

SEQ ID NO:23 (KLH2 Domäne g),

SEQ ID NO:24 (partielle KLH2 Domäne h);

- (b) Nukleinsäuresequenzen, die mit dem Gegenstrang einer Nukleinsäuresequenz nach (a) hybridisieren und für ein Polypeptid kodieren, das die immunologischen Eigenschaften wenigstens einer Domäne eines Hämocyanins aufweist;
- (c) Nukleinsäuresequenzen, die aufgrund des genetischen Codes zu den unter (a) und (b) definierten DNA-Sequenzen degeneriert sind und für ein Polypeptid kodieren, das die immunologischen Eigenschaften wenigstens einer Domäne eines Hämocyanins aufweist;
- (d) Nukleinsäuresequenzen, die mit einer der unter (a) bis (c) angegebenen Nukleinsäuresequenzen hybridisieren und deren Gegenstrang für ein Polypeptid kodiert, das die immunologischen Eigenschaften wenigstens einer Domäne eines Hämocyanins aufweist;
- (e) Nukleinsäuresequenzen, die wenigstens 60 % homolog zu einer der unter (a) angegebenen Nukleinsäuresequenzen sind;
- ~~(f) Varianten der unter (a) bis (d) angegebenen Sequenzen, wobei die Varianten gegenüber den unter (a) bis (d) angegebenen Sequenzen Additionen, Deletionen, Insertionen oder Inversionen aufweisen und für ein Polypeptid kodieren, das die immunologischen Eigenschaften wenigstens einer Domäne von Hämocyanin aufweist; und~~
- (g) Kombinationen mehrerer der unter (a) bis (f) angegebenen DNA-Sequenzen.

2. Nukleinsäuremolekül gemäß Anspruch 1, **dadurch gekennzeichnet**, daß die unter (b) oder (d) angegebene Hybridisierung unter stringenten Bedingungen durchgeführt wird.

3. Nukleinsäuremolekül gemäß Anspruch 1, **dadurch gekennzeichnet**, daß das unter (e) angegebene Nukleinsäuremolekül wenigstens 80 % homolog zu einer der unter (a) angegebenen Nukleinsäuresequenzen ist.

4. Nukleinsäuremolekül gemäß Anspruch 1, **dadurch gekennzeichnet**, daß das unter (e) angegebene Nukleinsäuremolekül wenigstens 90 % homolog zu einer der unter (a) angegebenen Nukleinsäuresequenzen ist.

5. Nukleinsäuremolekül gemäß Anspruch 1, **dadurch gekennzeichnet**, daß das unter (e) angegebene Nukleinsäuremolekül wenigstens 95 % homolog zu einer der unter (a) angegebenen Nukleinsäuresequenzen ist.

6. Nukleinsäuremolekül gemäß einem der Ansprüche 1 bis 5, **dadurch gekennzeichnet**, daß es ein Desoxyribonukleinsäuremolekül ist.

7. Konstrukt, umfassend ein Nukleinsäuremolekül gemäß einem der Ansprüche 1 bis 6.

8. Konstrukt gemäß Anspruch 7, weiterhin umfassend einen zur Expressionskontrolle geeigneten Promotor, wobei die für ein Hämocyanin, eine Hämocyanin-Domäne oder ein funktionelles Fragment davon kodierende Nukleinsäuresequenz unter der Kontrolle des Promotors steht.

9. Konstrukt gemäß Anspruch 7 oder 8, weiterhin umfassend eine für ein Antigen kodierende Nukleinsäuresequenz, die direkt mit der für ein Hämocyanin, eine Hämocyanin-Domäne oder ein funktionelles Fragment davon kodierenden Nukleinsäuresequenz verbunden ist.

10. Konstrukt gemäß Anspruch 9, wobei das Antigen ausgewählt ist aus: Tumorantigenen, Virusantigenen und Antigenen bakterieller oder parasitärer Pathogene.

11. Konstrukt gemäß einem der Ansprüche 7 bis 10, wobei das Konstrukt wenigstens einen Teil eines Vektors enthält, wobei der Vektor ausgewählt ist aus: Bakteriophagen, Adenoviren, Vacciniaviren, Baculoviren, SV40-Virus und Retroviren.

12. Konstrukt gemäß einem der Ansprüche 7 bis 11, wobei das Konstrukt weiterhin eine His-Tag-kodierende Nukleinsäuresequenz umfaßt und die Expression des Konstrukts zur Bildung eines Fusionsproteins mit einem His-Tag führt.

13. Wirtszelle, enthaltend ein Konstrukt gemäß einem der Ansprüche 7 bis 12, wobei die Wirtszelle eine zur Expression des Konstrukts geeignete prokaryontische oder eukaryontische Zelle ist.

14. Wirtszelle gemäß Anspruch 13, **dadurch gekennzeichnet**, daß die prokaryontische Wirtszelle ausgewählt ist aus E. coli und Bacillus subtilis.

15. Wirtszelle gemäß Anspruch 13, **dadurch gekennzeichnet**, daß die eukaryontische Wirtszelle ausgewählt ist aus Hefezellen, Insektenzellen und Säugerzellen, bevorzugt aus CHO-Zellen, COS-Zellen und HeLa-Zellen.

~~16. Verfahren zum Herstellen eines Hämocyanin-Polypeptides, wobei das Nukleinsäuremolekül gemäß einem der Ansprüche 1 bis 6 und/oder das Konstrukt gemäß einem der Ansprüche 7 bis 12 in einer geeigneten Wirtszelle exprimiert wird und das Protein gegebenenfalls isoliert wird.~~

17. Verfahren gemäß Anspruch 16, **dadurch gekennzeichnet**, daß das hergestellte Hämocyanin-Polypeptid natürlich oder chemisch modifiziert wird.

18. Verfahren gemäß Anspruch 17, **dadurch gekennzeichnet**, daß die Modifikation eine Quervernetzung oder eine kovalente Bindung an ein Antigen ist.

19. Verfahren gemäß einem der Ansprüche 16 bis 18, **dadurch gekennzeichnet**, daß die Expression in einer Wirtszelle gemäß einem der Ansprüche 13 bis 15 durchgeführt wird.

20. Hämocyanin-Polypeptid, umfassend eine Aminosäuresequenz, die von einem oder mehreren der Nukleinsäuremoleküle nach einem der Ansprüche 1 bis 6 kodiert wird.

21. Hämocyanin-Polypeptid gemäß Anspruch 20, umfassend wenigstens eine aus der folgenden Gruppe ausgewählte Aminosäuresequenz:

SEQ ID NO:25 (HtH1 Domäne a + Signalpeptid),

SEQ ID NO:26 (HtH1 Domäne b),

SEQ ID NO:27 (HtH1 Domäne c),

SEQ ID NO:28 (HtH1 Domäne d),

SEQ ID NO:29 (HtH1 Domäne e),

SEQ ID NO:30 (HtH1 Domäne f),

SEQ ID NO:31 (HtH1 Domäne g),

SEQ ID NO:32 (HtH1 Domäne h),

SEQ ID NO:33 (partielle HtH2 Domäne b),

SEQ ID NO:34 (HtH2 Domäne c),

SEQ ID NO:35 (HtH2 Domäne d),

SEQ ID NO:36 (HtH2 Domäne e),

SEQ ID NO:37 (HtH2 Domäne f),

SEQ ID NO:38 (HtH2 Domäne g),

SEQ ID NO:39 (HtH2 Domäne h),

SEQ ID NO:40 (partielle KLH1 Domäne b),

SEQ ID NO:41 (KLH1 Domäne c),

SEQ ID NO:42 (partielle KLH1 Domäne d),

SEQ ID NO:43 (partielle KLH1 Domäne e),

SEQ ID NO:44 (KLH2 Domäne b),

SEQ ID NO:45 (KLH2 Domäne c),

SEQ ID NO:46 (partielle KLH2 Domäne d),

SEQ ID NO:47 (KLH2 Domäne g),

SEQ ID NO:48 (partielle KLH2 Domäne h),

oder ein Fragment einer dieser Sequenzen, das die immunologischen Eigenschaften wenigstens einer Domäne eines Hämocyanins aufweist.

22. Rekombinantes Hämocyanin-Polypeptid, erhältlich durch das Verfahren gemäß einem der Ansprüche 16 bis 19 oder Modifikationen davon.

23. Rekombinantes Hämocyanin-Polypeptid gemäß Anspruch 22, **dadurch gekennzeichnet**, daß es die Sequenzen SEQ ID NO: 25 bis 32 umfaßt und Hämocyanin 1 aus *Haliotis tuberculata* ist.

24. Rekombinantes Hämocyanin-Polypeptid gemäß Anspruch 22, **dadurch gekennzeichnet**, daß es die Sequenzen SEQ ID NO: 33 bis 39 umfaßt und Hämocyanin 2 aus *Haliotis tuberculata* ist.

25. Rekombinantes Hämocyanin-Polypeptid gemäß Anspruch 23, **dadurch gekennzeichnet**, daß es ein scheinbares Molekulargewicht von 370 KDa in SDS-PAGE unter ~~reduzierenden Bedingungen aufweist.~~

26. Rekombinantes Hämocyanin-Polypeptid gemäß Anspruch 24, **dadurch gekennzeichnet**, daß es ein scheinbares Molekulargewicht von 370 KDa in SDS-PAGE unter reduzierenden Bedingungen aufweist.

27. Rekombinantes Hämocyanin-Polypeptid gemäß Anspruch 21, **dadurch gekennzeichnet**, daß das Hämocyanin-Polypeptid die Sequenzen SEQ ID NO: 40 bis 43 umfaßt und KLH1 aus *Megathura crenulata* ist.

28. Rekombinantes Hämocyanin-Polypeptid gemäß Anspruch 21, **dadurch gekennzeichnet**, daß das Hämocyanin-Polypeptid die Sequenzen SEQ ID NO: 44 bis 48 umfaßt und KLH2 aus *Megathura crenulata* ist.

29. Rekombinantes Hämocyanin-Polypeptid gemäß einem der Ansprüche 20 bis 28, **dadurch gekennzeichnet**, daß es kovalent an Viren, Virenbestandteile, Bakterien, Bakterienbestandteile, DNA, DNA-Bestandteile, anorganische oder organische Moleküle wie z. B. Kohlenhydrate Peptide und/oder Glykoproteine gebunden ist.

30. Rekombinantes Hämocyanin-Polypeptid gemäß einem der Ansprüche 20 bis 29, **dadurch gekennzeichnet**, daß das Hämocyanin-Polypeptid nicht-glykosyliert ist.

31. Rekombinantes Hämocyanin-Polypeptid gemäß einem der Ansprüche 20 bis 29, **dadurch gekennzeichnet**, daß das Hämocyanin-Polypeptid glykosyliert ist.

32. Pharmazeutische Zusammensetzung, enthaltend ein Nukleinsäuremolekül gemäß einem der Ansprüche 1 bis 6 und/oder ein Konstrukt gemäß einem der Ansprüche 7 bis 12 und physiologisch verträgliche Zusatzmittel.

33. Pharmazeutische Zusammensetzung gemäß Anspruch 32, **dadurch gekennzeichnet**, daß sie zur gentherapeutischen Behandlung von Tumoren verwendet wird.

~~34. Pharmazeutische Zusammensetzung, enthaltend ein Hämocyanin-Polypeptid nach einem der Ansprüche 20 bis 31 und physiologisch verträgliche Zusatzmittel.~~

35. Pharmazeutische Zusammensetzung gemäß Anspruch 34, **dadurch gekennzeichnet**, daß sie als Antiparasitenmittel, Antivirusmittel oder als Antitumormittel verwendet wird.

36. Pharmazeutische Zusammensetzung nach Anspruch 34, **dadurch gekennzeichnet**, daß sie zum Behandeln einer der folgenden Erkrankungen verwendet wird: Schistosomiasis, Bluthochdruck, Oberflächen-Harnblasenkarzinomen, Epithelkarzinomen, Ovarialkarzinom, Mammakarzinom, Bronchialkarzinom und Kolonrektalkarzinom.

37. Pharmazeutische Zusammensetzung nach Anspruch 34, **dadurch gekennzeichnet**, daß sie als Impfstoff verwendet wird.

38. Pharmazeutische Zusammensetzung gemäß Anspruch 34, **dadurch gekennzeichnet**, daß sie zur Kokain-Mißbrauchsvorsorge verwendet wird.

39. Verwendung von Hämocyanin-Polypeptid gemäß einem der Ansprüche 20 bis 31 als Trägerstoff für Arzneimittel.

40. Liposom, umfassend ein Nukleinsäuremolekül gemäß einem der Ansprüche 1 bis 6, ein Konstrukt gemäß einem der Ansprüche 7 bis 12 und/oder ein Hämocyanin-Polypeptid gemäß einem der Ansprüche 20 bis 31.

41. Liposom gemäß Anspruch 40, **dadurch gekennzeichnet**, daß das Liposom weiterhin Zellerkennungsmoleküle umfaßt.

42. Antikörper, erhältlich durch Immunisieren eines Versuchstieres mit dem rekombinanten Hämocyanin-Polypeptid gemäß einem der Ansprüche 20 bis 31.

43. Screening-Verfahren zum Identifizieren von Tumor-spezifischer DNA in einer Zelle umfassend:

a) das Inkontaktbringen zellulärer DNA und/oder zellulären Proteins mit einer Sonde umfassend die Nukleinsäuresequenz gemäß einem der Ansprüche 1 bis 6 und/oder des Antikörpers gemäß Anspruch 42 und

b) das Nachweisen der spezifischen Bindung.

44. Screening-Verfahren gemäß Anspruch 43, **dadurch gekennzeichnet**, daß der nachzuweisende Tumor Harnblasenkarzinom, Epithelialkarzinom, Ovarialkarzinom, Mammakarzinom, Bronchialkarzinom oder Kolonrektalkarzinom ist.

Zusammenfassung

Nukleinsäuremolekül, umfassend eine für ein Hämocyanin kodierende Nukleinsäuresequenz

Die vorliegende Erfindung betrifft ein Nukleinsäuremolekül, umfassend eine für ein Hämocyanin, eine Hämocyanin-Domäne oder ein Fragment mit den immunologischen Eigenschaften wenigstens einer Domäne von Hämocyanin kodierende Nukleinsäuresequenz.

Weiterhin betrifft die Erfindung Konstrukte, die das Nukleinsäuremolekül und gegebenenfalls einen zur Expressionskontrolle geeigneten Promotor enthalten. In einer bevorzugten Ausführungsform enthält das Konstrukt ferner eine für ein Antigen kodierende Nukleinsäuresequenz. Die Erfindung betrifft außerdem Wirtszellen, die diese Nukleinsäuremoleküle und/oder Konstrukte enthalten. Die Erfindung betrifft ferner die rekombinante Expression der Nukleinsäuremoleküle und/oder Konstrukte in den Wirtszellen. Weiterhin betrifft die Erfindung Hämocyanin, eine Hämocyanin-Domäne, ein Fragment mit den immunologischen Eigenschaften wenigstens einer Domäne von Hämocyanin und Hämocyanin-Fusionsproteine, die von den Nukleinsäuremolekülen und/oder Konstrukten kodiert werden. Die Erfindung betrifft weiterhin pharmazeutische Zusammen-

setzungen, die die Nukleinsäuremoleküle und/oder Hämocyanin, eine Hämocyanin-Domäne, ein Fragment davon oder ein Fusionsprotein, enthalten. Weiterhin betrifft die Erfindung Liposomen, die die Nukleinsäuremoleküle und/oder Hämocyanin, eine Hämocyanin-Domäne, ein Fragment davon oder ein Fusionsprotein, enthalten. Ferner betrifft die Erfindung Antikörper, die durch Immunisieren eines Versuchstieres mit dem Hämocyanin, einer Hämocyanin-Domäne, einem Fragment davon oder einem Fusionsprotein, erhältlich sind, und deren Verwendung in Screening-Verfahren zum Identifizieren von Tumoren.

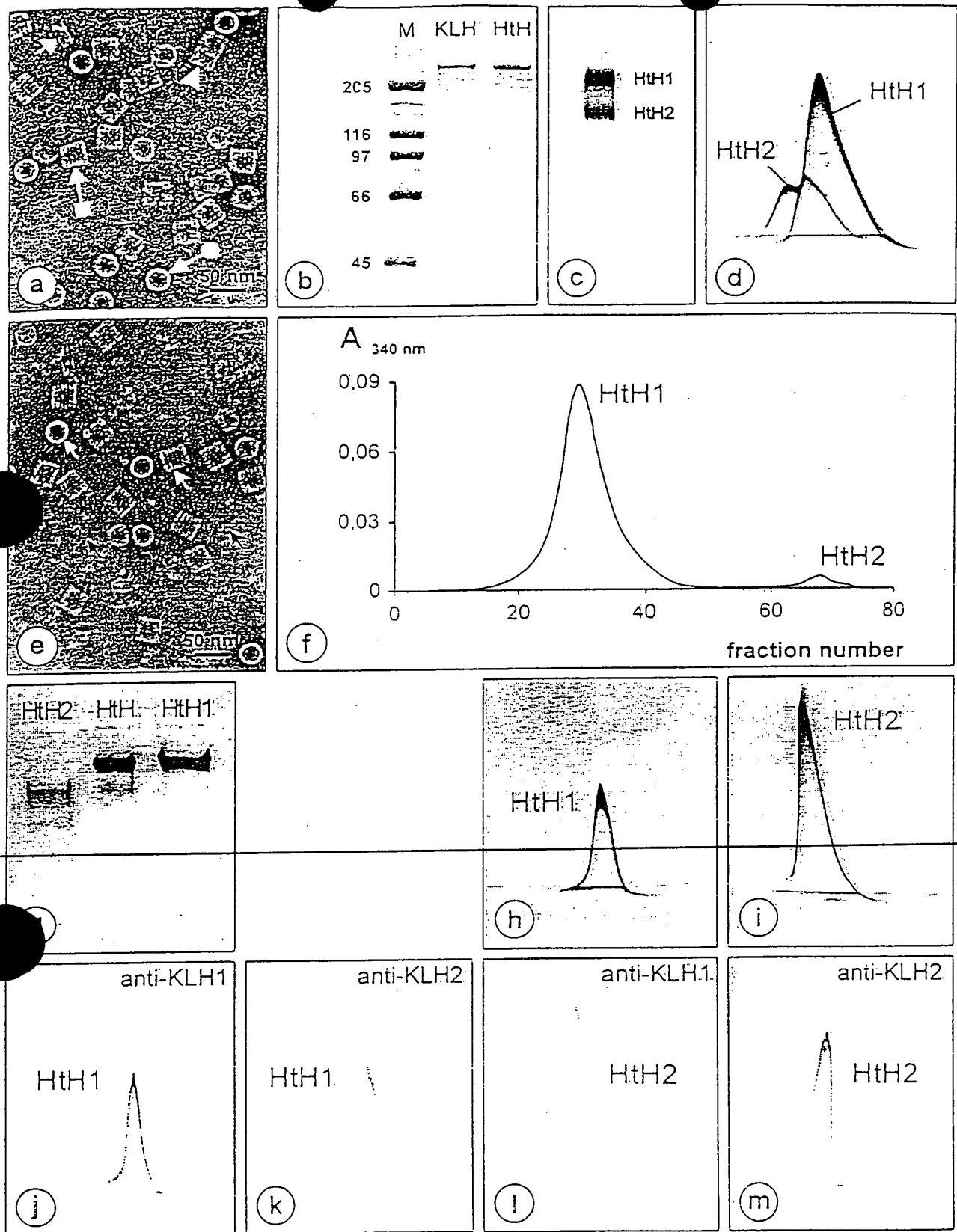


Fig. 1



Fig. 2

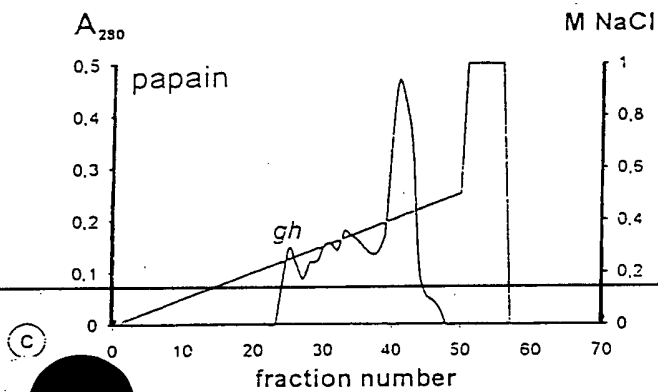
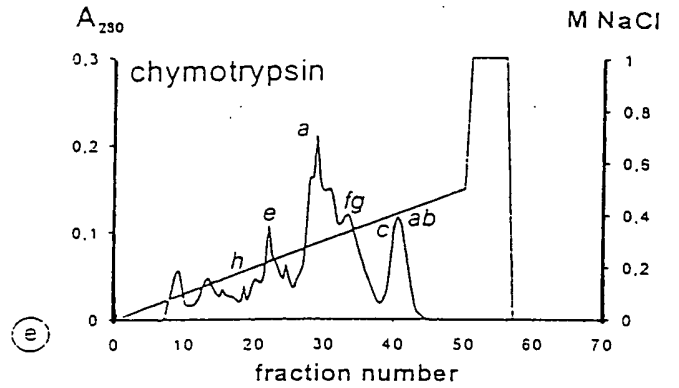
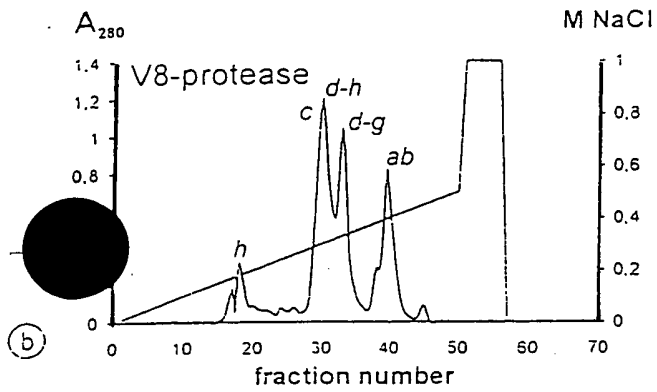
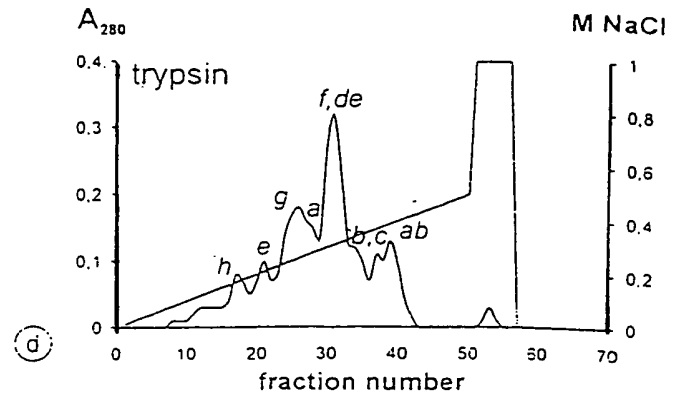
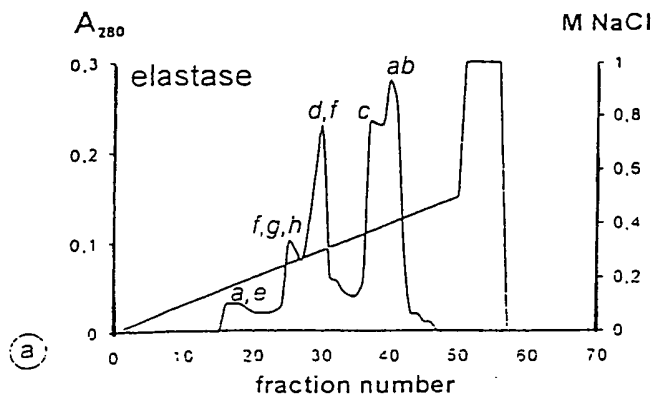


Fig. 3

Figur 4

cDNA-Sequenz in Verbindung mit Intronstruktur des HtH1

Domäne a

GGCTTGTTTCAGTTTCTACTCGTCGCCCTTGTGGCGGGGGCTGGAGCAGACAACGTCGTCAG
AAAGGACGTGAGTCACCTCACGGATGACGAGGTGCAAGCTCTCCACGGCGCCCTCCATGAC
GTCAGTGCATCTACAGGGCCTCTGAGTTTCTGAAGACATAACATCTTACCATGCCGCACCAG
CGTCGTGTGACTACAAGGGACGGAAGATCGCCTGCTGTGTCCACGGTATGCCCAGTTTCCC
CTTCTGGCACAGGGCATATGTCGTCCAAGCCGAGCGGGCACTGTTGTCCAAACGGAAGACT
GTCGGAATGCCTTACTGGGACTGGACGCAAACGCTGACTCACTTACCATCTCTTGTGACTG
AACCCATCTACATTGACAGTAAAGGTGGAAAGGCTCAAACCAACTACTGGTACCGCGGCGA
GATAGCGTTTCATCAATAAGAAGACTGCGCGAGCTGTAGATGATCGCCTATTCGAGAAGGTG
GAGCCTGGTCACTACACACATCTTATGGAGACTGTCTCGACGCTCTCGAACAGGACGAAT
TCTGTAAATTTGAAATCCAGTTCGAGTTGGCTCATAATGCTATCCATTACTTGGTTGGCGG
TAAATTTGAATATTCAATGTCAAACCTTGAATACACCTCCTACGACCCCATCTTCTTCCTC
CACCCTCCAACGTTGACCGCCTCTTCGCCATCTGGCAGCGTCTTCAGGAAGTGGCAGGAA
AGAATCCCAATGCAATGGACTGTGCACATGAACCTCGCTCACCAGCAACTCCAACCCTTCAA
CAGGGACAGCAATCCAGTCCAGCTCACAAAGGACCCTCGACACCTGCTGACCTCTTTGAT
TACAAACAACCTTGGATACAGCTACGACAGCTTAAACCTGAATGGAATGACGCCAGAACAGC
TGAAAACAGAACTAGACGAACGCCACTCCAAAGAACGTGCGTTTGCAAGCTTCGACTCAG
TGGCTTTGGGGGTTCTGCCAACGTTGTTGTCTATGCATGTGTCCCTGATGATGATCCACGC
AGTGATGACTACTGCGAGAAAGCAGGCGACTTCTTCATTCTTGGGGGTCAAAGCGAAATGC
CGTGGAGATTCTACAGACCCTTCTTCTATGATGTAACCTGAAGCGGTACATCACCTTGGAGT
CCCGCTAAGTGGCCACTACTATGTGAAAACAGAACTCTTCAGCGTGAATGGCACAGCACTT
TCACCTGATCTTCTTCCTCAACCAACTGTTGCCTACCGACCTGGGAAAG

Domäne b

GTCACCTTGACCCACCTGTGCATCATCGCCACGATGACGATCTTATTGTTTCGAAAAAATAT
AGATCATTGACTCGTGAAGAGGAATACGAGCTAAGGATGGCTCTGGAGAGATTCCAGGCC
GACACATCCGTTGATGGGTACCAGGCTACAGTAGAGTACCATGGCCTTCCTGCTCGTTGTC
CACCACCAGATGCAAAAGTCAGGTTGCGCTGTTGTATGCATGGCATGGCATCCTTCCTCA
CTGGCACCGGCTGTTGTTTACCCAGGTGGAAGATGCTCTTGTACGGCGTGGATCGCCTATC
GGTGTTCCTTATTGGGACTGGACAAAACCTATGACTCACCTTCCAGACTTGGCATCAAATG
AGACGTACGTAGACCCGTATGGACATACACATCATAATCCATTCTTCAATGCAAATATATC
TTTTGAGGAGGGACACCATCACACGAGCAGGATGATAGATTGAAACTGTTTGCCCCAGTC
GCTTTTGGGGAGCATTCCCATCTGTTTGTATGGAATCCTGTACGCATTTGAGCAGGAAGATT
TCTGCGACTTTGAGATTAGTTTGAAGTAGTCCATAATTCTATTTCATGCGTGGATAGGCGG
TTCCGAAGATTACTCCATGGCCACCCTGCATTACACAGCCTTTGACCCCATTTTCTACCTT
CATCATTCCAATGTCGATCGTCTATGGGCAATCTGGCAAGCTCTTCAAATCAGGAGACACA
AGCCATATCAAGCCCCTGTGCACAGTCTGTGGAACAGTTGCCAATGAAGCCATTGCTTT
CCCATCACCTCTTAACAACAACGAGAAGACACATAGTCATTTCAGTCCCGACTGACATTTAT
GACTACGAGGAAGTGTGCACTACAGCTACGATGATCTAACGTTTGGTGGGATGAACCTTG
AAGAAATAGAAGAAGCTATACATCTCAGACAACAGCATGAACGAGTCTTCGCGGGATTTCT
CCTTGCTGGAATAGGAACATCTGCACTTGTGACATTTTCATAAATAAACCGGGGAACCAA
CCACTCAAAGCTGGAGATATTGCCATTCTTGGTGGTGCCAAGGAAATGCCTTGGGCGTTTG
ACCGCTTGTATAAGGTGCAAAATACTGACTCATTGAAGACACTTTCTCTCGATGTGATGG
AGATTATGAAGTCACTTTTAAATTCATGATATGCACGGAACGCTCTTGATACGGACCTG
ATTCCACACGCAGCAGTTGTTTCTGAGCCAGCTCACC

Intron b/c

GTAAGTAAATTTACAAAATTTGGTGTCTCTAACTATCCTAAGTATTCAATCGTTAGCGTG
 TACCTATCTGCATAATGCAATACCCTGACTCCATATAAGTATAGTATATTTACTCTGGTGG
 AAAACAAACAAATTGAAAACAAGAGTGGACGTGCTGTTATGATTTCTTTTTCTATTCTTGGT
 TCGTTGTGTAATGCCACAGCCAGCAATTCCAGATATATAGCGACGGTCTATGAATACTCCA
 GTCTGGACCAGACAATCGTGTGGAATGGTTTAGGCACATTATATCAAATTCATTGTTGAAG
 ATATGAGTTATGAGGTCACAATGTTGTCTTGTACCCCGTGTGAGTAGTGACGTCATTTCA
 TGACTGAAATCTCTTCAACGCCGTTTAGCAATAATAGGCTCAGTAGTATTCAACCAATTAC
 AATCAGTAGAAAATCTCTATACTATTCTTATGTTGCATCCTGATATCCCTATGCAAAAAT
 TAGTCATCTAATAATAATCATTTTCGATAAAATACTTTGGGCAAACAAATCAATGTAACATCT
 ATTTTCTTTTCAG

Domäne c

CTACCTTTGAGGATGAAAAGCACAGCTTACGAATCAGAAAAAATGTCGACAGCTTGACTCC
 TGAAGAAACAAATGAACTGCGTAAAGCCCTGGAGCTTCTTGAAAATGATCATACTGCAGGT
 GGATTCAATCAGCTTGGCGCCTTCCATGGAGAGCCTAAATGGTGCCCTAATCCTGAAGCGG
 AGCACAAGGTTGCATGCTGTGTTTCATGGCATGGCTGTTTTCCCTCATTGGCACAGGCTTCT
 TGCTCTCCAGGCGGAGAATGCTCTTAGAAAGCATGGGTACAGTGGTGCTCTACCATACTGG
 GATTGGACTCGCCCCCTTTCCCAACTTCCTGATCTGGTTAGTCATGAGCAGTATACAGATC
 CTTCCGACCATCACGTGAAGCATAACCCGTGGTTCAATGGCCACATCGATACAGTAAATCA
 GGATACCACCAGAAGCGTACGGGAGGATCTTTATCAACAACCTGAATTTGGACATTTTCAG
 GATATTGCTCAACAAGTCCTCTTAGCATTAGAACAAGATGACTTCTGTTTCGTTTGAAGTGC
 AGTATGAGATTTCCCATAAATTTTATCCATGCACTTGTAGGAGGAACCGACGCTTATGGCAT
 GGCATCGCTGAGATATACAGCATACGATCCAATCTTTTTCTTGATCATTTCAAACACCGAC
 AGGATCTGGGCTATTTGGCAATCCCTGCAAAAATACAGAGGCAAACCGTACAACACTGCCA
 ACTGCGCCATAGAATCTATGAGAAGGCCCTGCAACCATTTGGACTAAGCAGTGCCATTAA
 CCCTGACAGAATCACCAGAGAGCATGCTATCCCGTTTGATGTCTTCAACTATAGAGATAAC
 CTTCAATTACGTATATGATACCCCTGGAATTTAATGGTTTGTGCGATTTTCAACTTGATAGAG
 AGCTGGAAAAAATCAAGAGTCACGAAAGAGTATTTGCTGGATTCTTGCTGTCGGGGATTAA
 AAAATCTGCTCTTGTGAAATTCGAAGTTTGTACTCCACCTGATAATTGTCATAAAGCAGGG
 GAGTTTTATCTACTCGGGGACGAAAACGAGATGGCTTGGGCCTATGACCGACTTTTCAAGT
 ATGATATTACTCAGGTTCTGGAAGCAAACCATCTACACTTCTATGATCATCTCTTCATTG
 CTACGAAGTCTTTGATCTTAAAGGAGTGAGTTTGGGAAGTACCTGTTCCACACTGCAAAT
 GTGGTACATGATTCCGGCACAG

Intron c/d

GTACGTGGATTTGATTACATAGCAATGCTATATGATTTTCAAGTAATTACAACCTCAAGTCAT
 GTAGCCGTTTTAGATTGCATTACATCAAACAGCATTGGATTAAATTGGGGGATTGTCCAGG
 CCGCATTATGTTGCATTCCGAAAATAGTTTGTGTCCAGTGTCCACGTTTAAAATTAAACCA
 TTTTAATCATATTAGGGATAATTTTAATAGATGTTATAGTGCTTTATTTTCATATTGTTACA
 GTGGACAGTCACCAAGGACATATTTTACTCTATAGATACACAAACACCAATTAAAACCCCTG
 CTTTGGAAAGTCTAACTTTTTCCCCACAG

Domäne d

GCACCCGTGATCGTGATAACTACGTTGAAGAAGTTACTGGGGCCAGTCATATCAGGAAGAA
 TTTGAACGACCTCAATACCGGAGAAATGGAAAGCCTTAGAGCTGCTTTCCTGCATATTTCAG
 GACGACGGAACATATGAATCTATTGCCAGTACCATGGCAAACAGGCAAATGTCAATTGA
 ATGATCATAATATTGCGTGTGTGTCCATGGTATGCCTACCTTCCCCCAGTGGCACAGACT
 GTATGTGGTTTCAAGGTGGAGAATGCTCTCCTAAACAGGGGATCTGGTGTGGCTGTTCTTAC

TGGGAGTGGACTGCTCCCATAGACCATCTACCTCATTTTCATTGATGATGCAACATACTTCA
 ATTCCCAGACAACAGCGGTACGACCCTAACCCCTTTCTTCAGGGGAAAGGTTACTTTTGAAAA
 CGCAGTCACAACAAGGGACCCACAAGCCGGGCTCTTCAACTCAGATTATATGTATGAGAAT
 GTTTTACTTGGACTGGAGCAGGAAAATTATTGTGACTTTGAAATTCAGTTTGAGCTTGTTC
 ATAACGCACTTCATTCCATGCTGGGAGGTAAAGGGCAGTACTCCATGTCCTCCCTGGACTA
 TTCTGCGTTTGATCCCGTCTTCTTCTACATCATGCCAACACGGACAGACTGTGGGCAATC
 TGGCAGGAACACTACAAAGATTCCGAGAAGTGCCTTATGAAGAAGCGAAGTGTGCAATCAACC
 TCATGTCATCAACCACTGAAGCCGTTTTCAGTGATCCACATGAGAATCACGACAATGTCACCTT
 GAAATACTCAAAACCAACAGGACGGATTTCGACTACCAGAACCACTTCGGATACAAGTATGAC
 AACCTTGAGTTCCATCACTTATCTATCCCAAGTCTTGATGCTACCCCTGAAGCAAAGGAGAA
 ATCACGACAGAGTGTTTGCGGGCTTCTTCTTTCATAACATAGGAACTTCTGCTGACATAAC
 TATCTACATATGTCTGCCTGACGGACGGCGTGGCAATGACTGCAGTCATGAGGCGGGAACA
 TTCTATATCCTCGGAGGCGAAACAGAGATGCCTTTTATCTTTGACCGTTTGTATAAATTTG
 AAATCACCAAACCACTGCAACAGTTAGGAGTCAAGCTGCATGGTGGAGTTTTCGAACTGGA
 GCTTGAGATCAAGGCATACAACGGTTCCTATCTGGATCCCCATACCTTTGATCCAACATC
 ATCTTTGAACCTGGAACAG

Intron d/e

GTAATGCCATCTTAATACAGTTCGTTTCGTTAAATTATATATGTTTCGTTTACAACACCATAC
 CTTGAATTGAGGTAATACATCACTTGATATTGATAATGTAATGGTAATTGTTCTTGTTTGT
 AAAACCGTTTCTGGGGTGTTTATTCATATCCACCTGGTGGATAGTGAGTAAACACATTCG
 GTTTAATATGGGTATCTAATGGACAGTGAAGTGTGCTGGCTAGGCAGATACCTTGTTTCT
 GTGAATGGAGGTAGTAGAAAGGGGTTTTGATGATTGCAG

Domäne e

ATACCCATATCTTGGACCACGACCATGAGGAAGAGATACTTGTGAGGAAGAATATAATTGA
 TTTGAGCCCAAGGGAGAGGGTTTCTCTAGTCAAAGCTTTGCAAAGAATGAAGAATGATCGC
 TCCGCTGATGGGTACCAAGCCATTGCCTCTTTCATGCCCTGCCACCACTCTGTCCCAATC
 CATCTGCAGCTCACCGTTATGCTTGCTGTGTCCATGGCATGGCTACATTTCCCCAGTGGCA
 CAGACTGTACACTGTTTCAGGTTTCAGGATGCCCTGAGGAGACATGGTTCACTTGTTGGTATT
 CCTTACTGGGACTGGACAAAACCACTCAACGAGTTACCCGAGCTTCTTTCTTCAGCAACAT
 TTTATCATCCAATCCGGAATATTAATATTTCAAATCCATTCTCGGGGCTGACATAGAATT
 TGAAGGACCGGGCGTTTCATACAGAGAGGCACATAAATACTGAGCGCCTGTTTCACAGTGGG
 GATCATGACGGATACCACAACCTGGTTCTTCGAAACTGTTCTCTTTGCTTTGGAACAGGAAG
 ATTACTGCGATTTTGAAATACAATTTGAGATAGCCCATAATGGCATCCACACATGGATTGG
 TGGAAGCGCAGTATATGGCATGGGACACCTTCACTATGCATCATATGATCCAATTTTCTAC
 ATCCACCATTACAGACGGACAGAATATGGGCTATTTGGCAAGAGCTGCAGAAGTACAGGG
 GTCTATCTGGTTTCGGAAGCAAACCTGTGCCATTGAACATATGAGAACACCCTTGAAGCCTTT
 CAGCTTTGGGCCACCCTACAATTTGAATAGTCATACGCAAGAATATTCAAAGCCTGAGGAC
 ACGTTTGACTATAAGAAGTTTGGATACAGATATGATAGTCTGGAATTGGAGGGGCGATCAA
 TTTCTCGCATTGATGAACTTATCCAGCAGAGACAGGAGAAAGACAGAAGCTTTGACAGGGTT
 CCTCCTTAAAGGTTTTGGTACATCCGCATCTGTGTCATTGCAAGTTTGCAGAGTTGATCAC
 ACCTGTAAAGATGCGGGCTATTTCACTATTCTGGGAGGATCAGCCGAAATGCCATGGGCAT
 TCGACAGGCTTTATAAGTATGACATTACTAAAACCTTTCACGACATGAACCTGAGGCACGA
 GGACACTTTCTCTATAGACGTAACATCACGTCTTACAATGGAACAGTACTCTCGGGAGAC
 CTCATTACAGACGCCCTCCATTATATTTGTACCTGGACGCC

Intron e/f

GTGAGTACCTGTTTGCCTAAGACTTCTGTAGGCTAAAAGTGAAGAAATATCAATTAATT
 TCAATTCACCCAACTTGAAAACGGTACCTATATAGGTTAACTTTTTGTCTACAGTAACT
 GAACATACCTACACATTTTCATGAAATGATCTCTCAATATTTTCCACCAACAG

Domäne f

ATAAACTCAACTCACGGAAACATACACCTAACAGAGTCCGCCATGAGCTAAGTAGCCTTAG
 TTCCCGTGACATAGCAAGCTTGAAGGCAGCTTTGACAAGCCTTCAACATGATAATGGGACT
 GATGGTTATCAAGCTATTGCTGCCTTCCATGGCGTTTCTGCGCAGTGCCACGAGCCATCTG
 GACGTGAG

Intron f(1)

GTAAATTTACAGAGCTTTATGAAGTGTGTTTACAGAGTGAAGAGACCAAGATATACTTATACC
 CAAAAGTAGCTAGCAACAGACGATTTCACTTGTTCGGACACTTTGTATTATACGTTGGAT
 CCCAAGGTAAACGGAAACGTAACCGAGAATCAGTCCGTAAAGTGAGTGAGTGGTGGG
 CTTAACGTCGCACTCAGCAATACCCAGCTATGTGGCGACTCTCAGATTTACTGCTGGAGG
 AGAACCTACATAGCCCGGTTTAACCCGTGTGGTATGTAGTAAGACCAGCGCGGCATGGCTG
 GTATCTGACGGACGAAGGGTGGCGCTGCACGTATTCAGTGGTACAACACTGCACCCCAAT
 TTCACCGACCGGAGAATGATCTCCCTTCGGAGATATCGCCTGCCTTCCACGGGATTCGA
 ACTCGGTGACCTTCAAGCCAGCGCGCTTCTAGCGGGGGCGATTAGAGGTNAAGGCCGACG
 GCTCTACCACCTTAACTATCCCCCGGCCCACTCCTGACGGAAATGTTTATAATTCAGCCT
 TTGTTTTCTTATTAACACTCTTGGCAGATTTTCTATAGATAATGGATTCACATGTAGACA
 GTCTCCCATTTGTTGTAAGTGTAGTCAAGAGTTAGAATCTGAATACATTCTCCAAGATGGA
 TCAAGGAAAACAATAATTACTTGATGTTGCAG

Domäne f(2)

ATCGCCTGTTGCATCCACGGCATGGCGACGTTTCTCACTGGCACCGGTTGTACACTCTGC
 AGTTGGAGCAAGCGCTGCGCAGACACGGGTCCAGTGTTGCTGTTCCATACTGGGACTGGAC
 CAAGCCAATCACCGAATGCCACACATTCTGACAGACGGAGAATATTATGACGTTTGGCAA
 AATGCCGTCTTGCCCAATCCGTTTGCAAGAGGTTATGTGAAAATTAAAGATGCATTTACGG
 TGAGAAATGTCCAGGAAAGTCTGTTCAAAATGTCAAGTTTTGAAAGCACTCGCTTCTGTT
 TGACCAGGCTTTGTTGGCTCTTGAACAACTGACTACTGTGACTTCGAAGTTCAGTTTGAA
 GTGATGCATAACACGATCCATTATCTCGTAGGAGGCGTCAAACGTACGCCTTCTCCTCTC
 TCGAGTATTCTCATACGATCCAATCTTCTTTATTCACCACTCGTTTGTGACAAAATATG
 GGCTGTATGGCAAGAACTGCAAGCAGGAGACATCTACAGTTTAGAACAGCTGATTGTGCT
 GTGGGCCTCATGGGTGAGGCAATGAGGCCTTTCAACAAGGATTTCACCACAACCTCGTTCA
 CCAAGAAGCACGCAGTCCCTAATACAGTATTTGATTATGAAGATCTTGGCTATAACTATGA
 CAACCTTGAAATCAGTGGTTTAAACTTAAATGAGATCGAGGCGTTAATAGCAAAACGCAAG
 TCACATGCTAGAGTCTTTGCTGGGTTTCTGTTGTTTGGATTAGGAACTTCGGCTGATATAC
 ATCTGGAAATTTGCAAGACATCGGAAAACGCCATGATGCTGCTGTTGATTTTCATCCTTGG
 AGGTTCTGCAGAGATGCATTGGGCATACAACCGCCTCTACAAGTATGACATTACAGAAGCA
 TTGCAGGAATTTGACATCAACCCTGAAGATGTTTTCCATGCTGATGAACCATTTTTCTCTGA
 GGCTGTGCGGTTGTTGCTGTGAATGGAAGTGTCAATCCATCGTCTCATCTTACCAGCCAAC
 GATAATCTATGAACCAGGCGAAG

Intron f/g

GTGAGATATATGCAAATTGAATGTTGTCCAGATGCGTTGTTTACATTTATATGCTTGGAAT
 TGTCTGAACGAATACAGTGGAATAACCAAAAGCTGAAAAATAAAAAGATATATACTTCAT
 TCTGAATTTGTGAGTATTGCTGACCCAAAAACAGTTATCCATGTCGACACTATATTTGCC
 TTTCTGAATCTGAGACTGCGTTATGTTTCTAATAATCACGAAATATGGTATACAGGTTGTG
 TATCTGTAGAATACCCAAGGCAGAATTTAAAGGGTCACACCCTGTTTAATACAG

Domäne g(1)

ATCACCATGACGACCATCAGTCGGGAAGCATAGCAGGATCCGGGGTCCGCAAGGACGTGAA
 CACCTTGACTAAGGCTGAGACCGACAACCTGAGGGAGGCGCTGTGGGGTGTTCATGGCAGAC
 CACGGTCCCAATGGCTTTCAAGCTATTGCTGCTTTCCATGGAAAACAGCTTTGTGTCCCA
 TGCCTGATGGCCACAACCTACTCATGTTGTACTCACGGCATGGCTACCTTCCCACACTGGCA
 TCGCCTCTACACCAAGCAGATGGAGGATGCAATGAGGGCGCATGGGTCTCATGTTCGGCCTG
 CCCTACTGGGACTGGACTGCTGCCTTCACCCACCTGCCAACACTGGTCACCGACACGGACA
 ACAACCCCTTCCAACAT

Intron g(2)

GTAAGAGCGGGGTAGGGATGGGGTGGTAGGGGGTGGGTTGTTCTATTACTTCCCGCTTCAC
 TTGTATGAAATGGATAACCTTGGCTGCATCCCAATTGCGTGATCGATTCTCTTTTCGATTCA
 CTCGTGCGATTAGACTGCCTTATTTACTATAGTAGTTAGAATGTTGCTCAGTGCGCCGTTA
 AACAACTAATACACAAAACCGCATTTGTTTTATATGGTCACTCTACTGTTTATCACGTATA
 TGTATGTTCCGACTCACTGGTTGGTGCGTACCATTCTACTGTCACACTGAGAGCCAATGTT
 CTCAGATGTGTGAAATGTTTGAAAGCCGTTTCTACATAATATTGCAGGAATACCATTGTAG
 AATGTAGTCAAACAGGTAACAATCTGTTAGTGAGCCAGTTCGAGGTTGCGTTGTAGGGTG
 TAGTCCAACAGGTAGGCAGTCCATAAGCATAGTTTTTAAGCATTTTAGATCATCTATAATT
 AACCACATGGTTAGCCGCTATGTTTAGTTAATCCAGTATAAGTTAGAACTGTTATATTTT
 GAAGGGAAGTGAGTAAATCCTTATTCCTTGACTACCATTTAATAGATTTCCCAATGACTCC
 ATTCAACTCCTAACTTTACATCACTGCTCTCTTCAACAG

Domäne g(2)

GGACACATTGATTATCTCAATGTCAGCACAACTCGATCTCCCCGAGACATGCTGTTCAACG
 ACCCCGAGCATGGATCAGAGTCGTTCTTCTACAGACAAGTCCTCTTAGCTCTGGAACAAAC
 TGATTTCTGCAAATTCGAAGTTCAGTTTGAGATAACCCACAATGCCATCCATTCCTGGACA
 GGTGGCCACAGCCCCCTACGGAATGTCCACTCTCGACTTCACTGCCTACGATCCTCTCTTCT
 GGCTTCACCACTCCAACACCGACAGAATCTGGGCTGTCTGGCAAGCTTTGCAAGAATACAG
 AGGACTTCCATACAACCATGCCAATTGTGAGATCCAGGCAATGAAAACGCCCTGAGGCCT
 TTCAGTGACGATATCAACCAACCCACTCACAACCCCTAACCGCAAGCCATTAGATGCTGT
 TCGAGTATAATCGGTTGAGCTTCCAGTACGACAACCTCATCTTCCATGGATACAGTATTCC
 GGAAC TTGATCGCGTGCTTGAAGAAAGAAAGGAGGAGGACAGAATATTTGCTGCCTTCCTT
 CTCAGTGGAATCAAGCGTAGTGCTGATGTAGTGTTTCGACATATGCCAGCCAGAACACGAAT
 GTGTGTTTCGAGGGACTTTTGCGATTTTGGGAGGGGAGCTAGAAATGCCCTGGTCCTTCGA
 CAGACTGTTCCGCTATGATATCACCAAGGTGATGAAGCAGCTACACCTGAGGCATGACTCT
 GACTTTACCTTCAGGGTGAAGATTGTCGGCACCGACGACCACGAGCTTCCTTCAGACAGTG
 TCAAAGCACCAACTATTGAATTTGAACCGGGCG

Intron g/h

GTGAGTACGACAGGCATTTCTAGTAAAAACCTACTTTTGGTAAAAGGTTTCGAGAAATCACT
 TGAAGCAACAACATGATTTTGTAACGCCTATTACACGTGAACATGTCACACCCGGTGATGC
 CGTTTAATGGACATGCCTCTGTTAATGAAAGGGGTAAGTACATGTGTATGGGGATGGGATG
 GGAGCCACCTGTCCCAATTCATAGGTCCCTAGGATCCCAGTTGCGTAGGAATCCCCTGAT
 TAATGCCTTGTTGAATTCCTCCTGGAATTGTCTTGCCCAAATTTTTACAAACCCGCCCGA
 TATACCTTGGAAATAATTGGGCCTAAGGGTGGGGCTTTTAAGGACCAAGAACCCAACTAA
 ACCCCAACCCATTTTTTCCCACCCATTCCAGGTTTTGTTTTACCAAATAAAAAGGTTTCCA
 CTTTGAGGAAACCCTTTAAGGGTCTTTTCAGGGCTTTTTTCTTTTCTGGGAATTCCAAT
 TCCGGGGGAACAAAATACATATATTTACAGACCTTTGGTCAAATTTATATAATTTCCGAC
 TTCATGTCATAGGTTTGTCTTTCTTCCTACACAG

Domäne h

TGCACAGAGGCGGAAACCACGAAGATGAACACCATGATGACAGACTCGCAGATGTCCTGAT
CAGGAAAGAAGTTGACTTCCTCTCCCTGCAAGAGGCCAACGCAATTAAGGATGCACTGTAC
AAGCTCCAGAATGACGACAGTAAAGGGGGCTTTGAGGCCATAGCTGGCTATCACGGGTATC
CTAATATGTGTCCAGAAAGAGGTACCGACAAGTATCCCTGCTGTGTCCACGGAATGCCCCG
GTTCCCCCACTGGCACCGCCTGCATACCATTGAGATGGAGAGAGCTCTGAAAAACCATGGC
TCTCCAATGGGCATTTCCTTACTGGGATTGGACAAAGAAGATGTCGAGTCTTCCATCTTTCT
TTGGAGATTCCAGCAACAACAACCTTTCTACAAATATTACATCCGGGGCGTGCAGCACGA
AACAAACCAGGGACATTAATCAGAGACTCTTAAATCAAACCAAGTTTGGTGAATTTGATTAC
CTATATTACCTAACTCTGCAAGTCCTGGAGGAAAACTCGTACTGTGACTTTGAAGTTCAGT
ATGAGATCCTCCATAACGCCGTCCTGCTGGCTTGGAGGAACTGGAAAGTATTCCATGTC
TACCCTGGAGCATTCGGCCTTTGACCCTGTCTTCATGATTCACTACTCGAGTTTGGATAGA
ATCTGGATCCTTTGGCAGAAGTTGCAAAAGATAAGAATGAAGCCTTACTACGCATTGGATT
GTGCTGGCGACAGACTTATGAAAGACCCCTGCATCCCTTCAACTACGAAACCGTTAATGA
AGATGAATTCACCCGCATCAACTCTTTCCCAAGCATACTGTTTGACCACTACAGGTTCAAC
TATGAATACGATAACATGAGAATCAGGGGTCAGGACATACATGAACTTGAAGAGGTAATTC
AGGAATTAAGAAACAAAGATCGCATATTTGCTGGTTTTGTTTTGTCGGGCTTACGGATATC
AGCTACAGTGAAAGTATTCATTCATTCGAAAAACGATACAAGTCACGAAGAATATGCAGGA
GAATTTGCAGTTTTTGGGAGGTGAGAAGGAGATGCCGTGGGCATATGAAAGAATGCTGAAAT
TGGACATCTCCGATGCTGTACACAAGCTTCACGTGAAAGATGAAGACATCCGTTTTAGAGT
GGTTGTTACTGCCTACAACGGTGACGTGTTACCACCAGGCTGTCTCAGCCATTTCATCGTC
CACCGTCCAGCCCATGTGGCTCACGACATCTTGGTAATCCCAGTAGGTGCGGGCCATGACC
TTCCGCCTAAAGTCGTAGTAAAGAGCGGCACCAAAGTCGAGTTTACACCAATAGATTCTGC
GGTGAACAAAGCAATGGTGGAGCTGGGCAGCTATACTGCTATGGCTAAATGCATCGTTCCC
CCTTTCTCTTACCACGGCTTTGAACTGGACAAAGTCTACAGCGTCGATCACGGAGACTACT
ACATTGCTGCGAGGTACCCACGCGTTGTGTGAGCAGAACCTCAGGCTCCACATCCACGTGGA
ACACGAGTAG

3' UTR

TTCACAG

Intron UTR

GTGAGGAGAAGGCCCCAGGCTAGCAGGGCAATGGATGAAGGAAATAGGGGCAAAGGGAATA
GCAGTTACACCATCGACATTTCCAACCTCCTCAGAACTAATATATAGCCTTAATACAACC
AGCCAAGACTCAACGGGCAGCCGGGGTGGGGGGATTGTTGGTTCGCTGTTTCAGACCAGGG
TGCAAAATATCAGTGCACAAATCAACATGTTGCGTGTGACACACTGACACAGCAGTCATTG
AACCTGCAGACCCATAACAGGAAAATGGGGCAGATACGATCAAAGACAGTGTAATATAGGG
ATAAGTAGGCATATGCAACCACCTGATGGAAATGAAAAGGGGTAAGTTTAAACCCCGGCTA
CCAAAGGTCCAATGGTTCCTTAACCCAGCTTACGCTATCCCTCTAATTTTCAAGTATTGAGCT
GATTTCTGTGAGTTTCATGTAACTGTATACTTTCTGTATTATTACAG

3' UTR

GTTGCTATGCCGACTGCGCTATATTGGTGAACGAGACGATGAGGACATCTCTGAAAGAGTT
CGCCAAGTGATGTGTAGGTCACGGAAGTATTGTTGAGCTAACAATATGATGATTTCAAAT
GACTTGGCGCTCTAGGACAAAGACATAATTCATCAGCACCCCTGTGCACCAACTCTTTGTTT
GCTGCAACGTCGTGACAAGCGACACGTCAATCAACAAGCTGTTCAAACCTCAAGTGGATGTA
ACTAGAATCGTTGGGCCATCGTTTCAAAAGTATTGACAGATGTCACACATGATGGCGAGAA
ACACTTTAGAATTTTAAATGACCTAGAGTGACTTGTAATATGTAAATATATTCTTCAAAG
ACTCAGCTGAACTATTGTTGGATAACACATCAATTCCCTCAACAAAATGCTTTATCTTAC
ATGGATGTATGTAATGTGGCCGGCAATAAAGTATATATATGTATAAAAAAAAAAAAAAAAAA
A

Abgeleitete Primärstruktur des HtH1

Signalpeptid

LVQFLLVALVAGAGA

Domäne a

DNVVRKDVSHLTDDEVQALHGALHDVTASTGPLSFEDITSYHAAPASCDYKGRKIACCVHG
MPSFPFWHRAYVVQAERALLSKRKTVGMPYWDWTQTLTHLPSLVTEPIYIDSKGGKAQTNY
WYRGEIAFINKKTARAVDDRLFEKVEPGHYTHLMETVLDALDAQDEFCKFEIQFELAHNAIH
YLVGGKFEYSMSNLEYTSYDPIFFLHHSNVDRLFAIWQRLQELRGKNPNAMDCAHDLAHQQ
LQPFNRDSNPVQLTKDHSTPADLFDYKQLGYSYDSLNLNGMTPEQLKTELDERHRSKERAF
SFRLSGFGGSANVVVYACVPDDDPRSDDYCEKAGDFFILGGQSEMPWRFYRPFYDVTEAV
HHLGVPLSGHYVKTFLFSVNGTALSPDLLPQPTVAYRPGK

Domäne b

GHLDPVHHRHDDDLIVRKNIDHLTREEEYELRMALERFQADTSVDGYQATVEYHGLPARC
PRPDAKVRFACCMHGMASFPHWHRLFTQVEDALVRRGSPIGVPYWDWTKPMTHLPDLAS
ETYVDPYGHTHHNPFFNANISFEEGHHHTSRMIDSKLFAPVAFGEHSHLFDGILYAFEQED
FCDFEIQFELVHNSIHAWIGGSEDYSMATLHYTAFDPIFYLHHSNVDRLWAIWQALQIRRH
KPYQAHCAQSVEQLPMKPFAFSPPLNNNEKTHSHSVPTDIYDYEEVLHYSYDDLTFGGMNL
EEIEEAHILRQQHERVFAGFLLAGIGTSALVDIFINKPGNQPLKAGDIAILGGAKEMPWAF
DRLYKVEITDSLKTLSDVDGDYEVTFKIHDMHGNAIDTDLI PHAAVVSEPAH

Domäne c

PTFEDEKHSRLIRKNVDSLTPREETNELRKALELLENDHTAGGFNQLGAFHGEPKWCPNPEA
EHKVACCVHGMVFPWHRLALQAENALRKHGYSALPYWDWTRPLSQLPDLVSHEQYTD
PSDHHVKHNPWFNGHIDTVNQDTTRSVREDLYQQPEFGHFTDIAQQVLLALEQDDFCSFEV
QYEISHNFIHALVGGTDAYGMASLRYTAYDPIFFLHHSNTDRIWAIWQSLQKYRGKPYNTA
NCAIESMRRLPLQPFGLSSAINPDRI TREHAIPFDVFNRYRDNLHYVYDTLEFNGLSISQLDR
ELEKIKSHERVFAGFLLSGIKKSALVKFEVCTPPDNCHKAGEFYLLGDENEMAWAYDRLFK
YDITQVLEANHLHFYDHLFIRYEVFDLKGVS LGTDLFHTANVVHDSGT

Domäne d

GTRDRDNYVEEVTGASHIRKNLNDLNTGEMESLRAAFLHIQDDGTYESIAQYHGKPGKCQL
NDHNIACCVHGMPTFPQWHRLYVVQVENALLNRGSGVAVPYWEWTAPIDHLPHFIDDATYF
NSRQQRYDPNPFGRKVTFENAVTTRDPQAGLFNSDYMENVLLALEQENYCDFEIQFELV
HNALHSM LGGKGQYSMSLDYSAFDPVFFLHHANTDRLWAIWQELQRFRELPYEEANCAIN
LMHQPLKPFSDPHENHDNVTLKYSKPQDGFQYQNHFGYKYDNLEFHLSIPSLDATLKQRR
NHDRVFAFGLLHNIGTSADITIIYICLPDGRRGNDCSHEAGTFYILGGETEMPFI FDRLYKF
EITKPLQQLGVKLHGGVFEELEIKAYNGSYLDPHTFDPTIIFEPGT

Domäne e

DTHILDHDHEEEILVRKNIIDLSPRERVS LVKALQRMKNDRSADGYQAIASFHALPPLCPN
 PSAAHRYACCVHGMATFPQWHRLYTVQVQDALRRHGSLVGIPYWDWTKPVNELPELLSSAT
 FYHPIRNINISNPFLGADIEFEGPGVHTERHINTERLFHSGDHDGYHNWFFETVLFALQE
 DYCDFEIQFEIAHNGIHTWIGGS AVYGMGHLHYASYDPIFYIHHSQTDRIWAIWQELQKYR
 GLSGSEANCAIEHMRTPLKPFSGPPYNLSHTOEYSKPEDTFDYKKFGYRYDSLELEGRS
 ISRIDELIQORQEKDRTFAGFLLKGFGTSASVSLQVCRVDHTCKDAGYFTILGGS AEMPWA
 FDRLYKYDITKTLHDMNLRHEDTFSIDVTITSYNGTVLSGDLIQTPSIIFVPGR

Domäne f

HKLNSRKHTPNRVRHELSSLSSRDIASLKAALTSLOHDNGTDGYQAIAAFHGVPAQCHEPS
 GREIACCIHGMATFPHWHRLYTLQLEQALRRHGSSVAVPYWDWTKPITELPHILTDGEYYD
 VWQNAVLANPFFARGYVKIKDAFTVRNVQESLFKMSSFGKHSLLFDQALLALEQTDYCDFEV
 QFEVMHNTIHYLVGGRQTYAFSSLEYSSYDPIFFIHHSFVDKIWAVWQELQSRRLQFRTA
 DCAVGLMGQAMRPFNKDFNHNSFTKKHAVPNTVFDYEDLGYNNDNLEISGLNLNEIEALIA
 KRKSHARVFAGFLLFGLGTSADIHLEICKTSENCHDAGVIFILGGS AEMHWAYNRLYKYDI
 TEALQEFDINPEDVFHADEPFFLRLSVVAVNGTVIPSSHLHQPTIIYEPGE

Domäne g

DHHDDHQSGSIAGSGVRKDVNTLTKAETDNLREALWGVMA DHGPNGFQAIAAFHGKPALCP
 MPDGHNYSCCTHGMATFPHWHRLYTKQMEDAMRAHGS HVGLPYWDWTAAFTHLPTLVTDTD
 NNPFOHGHIDYLVNSTTRSPRDMFLFNDPEHGSESFFYRQVLLALEQTD FCKFEVQFEITHN
 AIHSWTGGHSPYGMSTLDFTAYDPLFWLHHSNTDRIWAVWQALQEYRGLPYNHANCEIQAM
 KTPLRPFSD DINHNVPVTKANAKPLDVFEYNRLSFQYDNLI FHGYSIPELDRVLEERKEEDR
 IFAAFLLSGIKRSADVFDICQPEHECVFAGTFAILGGELEMPWSFDRLFRYDITKVMKQL
 HLRHDSDFTRVKIVGTDDHELPSDSVKAPTIEFEPG

Domäne h

VHRGGNHEDEHDDRLADVLRKEVDFLSLQEANAIKDALYKLQND DSKGGFEAIAGYHGY
 PNMCPERGTDKYPCCVHGMFPVFPWHRLHTIQMERALKNHGS PMGIPYWDWTKKMSSLPSF
 FGDSSNNNPYKYIIRGVQHETTRDVNQRLEFNQTKFGEFDYLYYLT LQVLEENSYCDFEVQ
 YEILHNAVHSWLGGTGQYSMTLEYSAFDPVFMIIHSSLDRIWILWQKLQKIRMKPYYALD
 CAGDRLMKDPLHPFNYETVNEDEFTRINSFPSILFDHYRFNYEYDNMRIRGQDIHELEEV
 QELRNKDRI FAGFVLSGLRISATVKVFIHKNDSHEEYAGEFAVLGGEKEMPWAYERMLK
 LDISDAVHKLHVKDEDIRFRVVVTAYNGDVVTTRLSQPFIVHRPAHVAHDILVIPVGAGHD
 LPPKVVKSGTKVEFTPIDSSVNKAMVELGSYTAMAKCIVPPFSYHGFELDKVYSVDHGDY
 YIAAGTHALCEQNLRLHIHVEHE

Figur 6

cDNA-Sequenz in Verbindung mit Intronstruktur des HtH2

Domäne b

CACAGACTGTTTCGTACCCAGGTGGAAGATGCTCTGATCAGGCGAGGATCGCCTATAGGGG
 TCCCCTACTGGGACTGGACTCAGCCTATGGCGCATCTCCAGGACTTGCAGACAACGCCAC
 CTATAGAGATCCCATCAGCGGGGACAGCAGACACAACCCCTTCCACGATGTTGAAGTTGCC
 TTTGAAAATGGACGTACAGAACGTCACCCAGATAGTAGATTGTTTGAACAACCTTTATTTG
 GCAAACATACGCGTCTCTTCGACAGTATAGTCTATGCTTTTGAAGCAGGAGGACTTCTGCGA
 TTTTGAAGTTCAATTTGAGATGACCCATAATAATATTCACGCCTGGATTGGTGGCGGCGAG
 AAGTATTCATGTCTTCTCTACACTACACAGCCTTCGACCCTATCTTCTACCTTCGTCCT
 CCAACACTGACCGGCTCTGGGCAATTTGGCAAGCGTTGCAGATACGAAGAAACAGGCCTTA
 CAAGGCTCATTGTGCTTGGTCTGAGGAACGCCAGCCTCTCAAACCTTTCGCCTTCAGTTCC
 CCACTGAACAACAACGAAAAACCTACGAAAACCTCGGTGCCCACCAACGTTTACGACTACG
 AAGGAGTCCTTGGCTATACTTATGATGACCTCAACTTCGGGGGCATGGACCTGGGTGAGCT
 TGAGGAATACATCCAGAGGCAGAGACAGAGAGACAGGACCTTGTGCTGGTTTCTTCTGTCA
 CATATTGGTACATCAGCGAATGTTGAAATCATTATAGACCATGGGACTCTTCATACCTCCG
 TGGGCACGTTTGTCTGTTCTTGGCGGAGAGAAGGAGATGAAATGGGGATTTGACCGTTTGTA
 CAAATATGAGATTACAGATGAAGTGAAGCAACTTAATCTCCGTGCTGATGATGTTTTTCAGC
 ATCTCTGTTAAAGTAACTGATGTTGATGGCAGTGAGCTGTCTCTGAACTCATCCCATCTG
 CTGCTATCATCTTCGAACGAAGCCATA

Intron b/c

GTAAGTAGCTACCTGTTTATTCAATTTTTTCGCTTTGCCAATCAATTCATTCAGCTTGAAA
 TTCAATAATTGTGTTTTGCATGGCTGAAAACCAATTTGAACTCTTTTCTTTCTCAGGTGCG
 AACTCAAATAAATAATCACTAATTGTTATGCACGCGGGTAGGGCATACTATATCCAC
 ATCGGTCACTCTCAAATGCAAACAATTGTCTTATTTCCGTTGGGACAAGCAAACCCCTT
 TCCTGTAATCTTGCTTTGGCATCCACTGGAATTAATGTTGACTGGTAATTGATACTGGCT
 CTCTTCTTGATAGAGTTAATATCTATAGTTTGTAATCTTTATGATTTTGCTATTTATAT
 TTCCAGACGATGCTATAGACACCTAGACTATTGTATAGCCACTTGTATTGTTTTCCATT
 TATTATTTATAACAGAACATGGCTTGTAATTTTTATTACCTTCCAG

Domäne c

TTGACCATCAGGACCCGCATCATGACACAATCATTAGGAAAAATGTTGATAATCTTACACC
 CGAGGAAATTAATTCTCTGAGGCGGGCAATGGCAGACCTTCAATCAGACAAAACCGCCGGT
 GGATTCCAGCAAATTGCTGCTTTTACGCGGGAACCCAAATGGTGCCCAAGTCCCGATGCTG
 AGAAGAAGTTCTCCTGCTGTGTCCATGGAATGGCTGTCTCCCTCACTGGCACAGACTCCT
 GACCGTGCAAGGCGAGAATGCCCTGAGAAAGCATGGATGTCTCGGAGCTCTCCCCTACTGG
 GACTGGACTCGGCCCCTGTCTCACCTACCTGATTTGGTTTTGGTAAGTAGCAGAACTACAC
 CGATGCCATATTCCACCGTGGAAGCCCGAAACCCCTGGTACAGCGGCCATATTGATACAGT
 TGGTGTTGACACAACAAGAAGCGTCCGTCAAGAAGTGTATGAAGCTCCTGGATTGGCCAT
 TATACTGGGGTCGCTAAGCAAGTGCTTCTGGCTTTGGAGCAGGATGACTTCTGTGATTTTG
 AAGTCCAGTTTGAGATAGCTCACAATTTTATTCACGCTCTTGTGCGGCGGAAGCGAGCCATA
 TGGTATGGCGTCACTCCGTTACACTACTTATGATCCAATTTTCTACCTCCATCATTCTAAC
 ACTGACAGACTCTGGGCTATATGGCAGGCTCTACAAAAGTACAGGGGCAAACCTTACAATT
 CCGCCAACTGCGCCATTGCTTCTATGAGAAAACCCCTACAACCCTTTGGTCTGACTGATGA
 GATCAACCCGGATGATGAGACAAGACAGCATGCTGTTCCCTTTCAGTGTCTTTGATTACAAG
 AACAACTTCAATTATGAATATGACACCCTTGACTTCAACGGACTATCAATCTCCAGCTGG
 ACCGTGAACTGTCACGGAGAAAGTCTCATGACAGAGTATTTGCCGGATTTTGTGCTGATGG

13 / 26

TATTCAGCAGTCTGCACTAGTTAAATTCTTTGTCTGCAAATCAGATGATGACTGTGACCAC
TATGCTGGTGAATTCTACATCCTTGGTGATGAAGCTGAAATGCCATGGGGCTATGATCGTC
TTTACAAATATGAGATCACTGAGCAGCTCAATGCCCTGGATCTACACATCGGAGATAGATT
CTTCATCAGATACGAAGCGTTTGATCTTCATGGTACAAGTCTTGGAAGCAACATCTTCCCC
AAACCTTCTGTACATACATGACGAAGGGGCAG

Intron c/d

GTGAGAACATTGATAATAGTTCAAATgAAGTATATCCGATTCAAGCTGTCGATACAAGATg
AGATACATAATCACAATGTTTGTATTAGATATCTCTCTTAATTTAATGCCGCTTTTATCAA
TATTCGAGCAATCCTTCAGCAACATACACCAGCAAATGTTTCATCAACAGACTATATTATT
TAATCTTTTAAAAATCCTTTTCTGTTGTTATAAACTTAAAGTATCGAATTCCTTGAATG
CGTCTTCTCTGCAGCATATAGTTAAGTTGTTGTGTTTCTCTGTGACG

Domäne d

GTCACCATCAGGCTGACGAGTACGACGAAGTTGTAAGTCTGCAAGCCACATCAGAAAGAA
TTTAAAAGATCTGTCAAAGGGGAGAAGTAGAGAGCCTAAGGTCTGCCTTCCTGCAACTTCAG
AACGACGGAGTCTATGAGAATATTGCCAAGTTCCACGGCAAGCCTGGGTTGTGTGATGATA
ACGGTCGCAAGGTTGCCTGTTGTGTCCATGGAATGCCACCTTCCCCCAGTGGGCACAGGCT
CTATGTCTCCAGGTGGAGAATGCTTTGCTGGAGAGAGGATCTGCCGTCTCTGTGCCATAC
TGGGACTGGACTGAAACATTTACAGAGCTGCCATCTTTGATTGCTGAGGCTACCTATTTCA
ATTCCCGTCAACAAACGTTTGACCCTAATCCTTTCTTCAGAGGTAAAATCAGTTTTGAGAA
TGCTGTTACAACACGTGATCCCCAGCCTGAGCTGTACGTTAACAGGTACTACTACCAAAC
GTCATGTTGGTTTTTGAACAGGACAACACTACTGCGACTTCGAGATACAGTTTGAGATGGTTC
ACAATGTTCTCCATGCTTGGCTTGGTGGAAGAGCTACTTATTCTATTTCTCTCTTGATTA
TTCTGCATTGACCCCTGTGTTTTCTTCACCATGCGAACACAGATAGATTGTGGGCCATC
TGGCAGGAGCTGCAGAGGTACAGGAAGAAGCCATACAATGAAGCGGATTGTGCCATTAACC
TAATGCGCAAACCTCTACATCCCTTCGACAACAGTGATCTCAATCATGATCCTGTAACCTT
TAAATACTCAAAACCCACTGATGGCTTTGACTACCAGAACAACCTTTGGATACAAGTATGAC
AACCTTGAGTTCAATCATTTCAGTATTCCCAGGCTTGAAGAAATCATTTCGtATTAGACAAC
GTCAAGATCGTGTGTTTGCAGGATTCCTCCTTCACAACATTGGGACATCCGCAACTGTTGA
GATATTCTGTCTGTGTCCCTACCACCAGCGGTGAGCAAACTGTGAAAACAAAGCCGGAACA
TTTGGCGTACTCGGAGGAGAAACAGAGATGGCGTTTCATTTTGACAGACTCTACAGGTTTC
ACATCAGTGAAACACTGAGGGACCTCGGCATACAGCTGGACAGCCATGACTTTGACCTCAG
CATCAAGATTCAAGGAGTAAATGGATCCTACCTTGATCCACACATCCTGCCAGAGCCATCC
TTGATTTTTGTGCCTGGTTCAAGT

Intron d/e

AAGAAAGTTTCACTGTCTAAATCTTTTTTTATGATAGAGGGTAGAGAAGTGGAGACAATGT
GACAAATATATTGAATAAAGTTGTTTAAAATTTATAACTCTCATAAGTTCATATTATGCTGA
AGCTGTAGCCATCTATAACTGTGTAACATGAAATGTTAAGACATTAACCTAAATACTTCAG
CTGATAACAAAACAATGTTAATACATACGTCAATGTAACATTTTCTTATCTTTAGGTTATA
GCATAAACACTTCAGAGATACAGTGACGAAAACCTCTATTAAATATTTACAGGT

Domäne e

TCTTTCCTGCGTCCTGATGGGCATTGAGATGACATCCTTGTGAGAAAAGAAGTGAACAGCC
TGACAACCAGGGAGACTGCATCTCTGATCCATGCTCTGAAAAGTATGCAGGAAGACCATTC
ACCTGACGGGTTCCAAGCCATTGCCTCTTTCCATGCTCTGCCACCACTCTGCCCTTCACCA
TCTGCAGCTCACCGTTATGCTTGCTGTGTCCACGGCATGGCTACATTTCCCCAGTGGCACA
GATTGTACACTGTACAGTTCCAGGATGCACTGAGGAGACATGGAGCTACGGTAGGTGTACC
GTATTGGGATTGGCTGCGACCGCAGTCTCACCTACCAGAGCTTGTACCATGGAGACATAC

CATGATATTTGGAGTAACAGAGATTTCCCCAATCCTTTCTACCAAGCCAATATTGAGTTTG
 AAGGAGAAAACATTACAACAGAGAGAGAAGTCATTGCAGACAACTTTTTGTCAAAGGTGG
 ACACGTTTTTTGATAAACTGGTTCTTCAAACAAGCCATCCTAGCGCTGAGCAGGAAAACCTAC
 TGTGACTTTGAGATTCAGTTTGAAATTCTTCACAACGGCGTTACACGTGGGTTCGGAGGCA
 GTCGTACCTACTCTATCGGACATCTTCATTACGCATTCTACGACCCCTCTTTTCTACCTTCA
 CCATTTCCAGACAGACCGTATTTGgGCAATCTGGCAAGAACTCCAGGAACAGAGAGGGCTC
 TCGGGTGATGAGGCTCACTGTGCTCTCGAGCAAATGAGAGAACCATTGAAGCCTTTCAGCT
 TCGGCGCTCCTTATAACTGGAATCAGCTCACACAGGATTTCTCCCGACCCGAGGACACCTT
 CGACTACAGGAAGTTTGGTTATGAATATGACAATTTAGAATTCCTGGGAATGTCAGTTGCT
 GAACTGGATCAATACATTATTGAACATCAAGAAAATGATAGAGTATTTCGCTGGGTTCTCTGT
 TGAGTGGATTTCGGAGGTTCCGCATCAGTTAATTTCCAGGTTTGTAGAGCTGATTCACATG
 TCAGGATGCTGGGTACTTCACCGTTCTTGGTGGCAGTGCTGAGATGGCGTGGGCATTTGAC
 AGGCTTTACAAATATGACATTACTGAACTCTGGAGAAAATGCACCTTCGATATGATGATG
 ACTTCACAATCTCTGTCAGTCTGACCGCCAACAACGGAACCTGTCTGAGCAGCAGTCTAAT
 CCCAACACCGAGTGTCTATTTCCAGCGGGGACATC

Intron e/f

AAGTAGTAAACTGCTCAGATTGTTTTTCATAATTACTCCACTATTAAGTAAAAAGTACTAGT
 AATTCATAGTACTGTTACAGAGAAATGTAACACAATAGACCACAGAGTCCATTTGTTAA
 ACGCCTTTGGCTTGGTAAGTCTGAGGTTTTGGTGACTGATGGAAAGCTAAAATATATTTTG
 ACAG

Domäne f(1)

GTGACATAAATACCAGGAGCATGTCACCGAACCGTGTTTCGCCGTGAGCTGAGCGATCTGTC
 TGCGAGGGGACCTGTCTAGTCTCAAGTCTGCTCTGCGAGACCTACAGGAGGATGATGGCCCC
 AACGGATACCAGGCTCTTGCAGCCTTCCATGGGCTACCAGCAGGCTGCCATGATAGCCGGG
 GAAATGAGAT

Intron f

ATATTTAAAGTATTTTATCTTACGCATGACCCTGACCCTATTATTTTTTTAATCCTATGAT
 GAAACATTTTACTTAGACTGGCTTGTGAGCCCCAGGCAAAATGCAGTGTAAAAATACACTGA
 CAGAGGATTAGGCATTCTTGGGAGTACTGTATAGTTAGTTGCATACATATTAGCGTTCCCT
 CACTAAAACGAATCTCTGAATGCTATCAATTAAAGATCATGATGCTTTGATTGTGTCTACT
 GTATTTAAATGGTGTAAAGATTTGCAATTACAATATACACAAACACGTTTCTCTGCATCTC
 GGAGAATGCAATCTTTCGTTGTACGCGTCTGTTTTTCATATTTTATGCATGTAGTTTGCAC
 TACTTAGCGTCCAATAAATCCATTCACAAAATCACACAAACAAACGATTTTAGGAATGTGA
 CTGTAGCTGCAACGAATATACCTGATCCTTTCTTGTTCAGAT

Domäne f(2)

CGCATGTTGCATTCACGGGATGCCGACCTTCCCCAGTGGCACAGACTGTACACCCTGCAG
 TTGGAGATGGCTCTGAGGAGACATGGATCATCTGTGCCATCCCCTACTGGGACTGGACAA
 AGCCTATCTCCGAACTCCCCCTCGCTCTTACCAGCCCTGAGTATTATGACCCATGGCATGA
 TGCTGTGGTAAACAACCCATTCTCCAAAGGTTTTGTCAAATTTGCAAATACCTACACAGTA
 AGAGACCCACAGGAGATGCTGTTCCAGCTTTGTGAACATGGAGAGTCAATCCTCTATGAGC
 AAATCTTCTTGCTCTTGAGCAAACCGACTACTGTGATTTTGGAGGTACAGTTTGGAGTCTT
 CCATAACGTGATCCACTACCTTGTTGGTGGACGTCAGACCTACGCATTGTCTTCTCTGCAT
 TATGCCTCCTACGACCCATTCTTCTTTATACACCATTCTTTTGTGGATAAGATGTGGGTAG
 TATGGCAAGCTCTTCAAAGAGAGGAGGAACTTCCATACAAGCGAGCTGACTGTgctGTCAA
 CCTAATGACTAAACCAATGAGGCCATTTGACTCCGATATGAATCAGAACCATTACAAAG
 ATGCACGCAGTTCCCAACACACTCTATGACTACGAGACACTGTACTACAGCTACGATAATC
 TCGAAATAGGTGGCAGGAATCTCGACCAGCTTCAGGCTGAAATTGACAGAAGCAGAAGCCA

CGATCGCGTTTTTGGCTGGATTCTTGCTTCGTGGAATCGGAACCTTCTGCTGATGTCAGGTTT
 TGGATTTGTAGAAATGAAAATGACTGCCACAGGGGTGGAATAATTTTCATCTTAGGTGGAG
 CCAAGGAAATGCCATGGTCATTTGACAGAACTTCAAGTTTGATATCACCCATGTACTCGA
 GAATGCTGGCATTAGCCCAGAGGACGTGTTTGATGCTGAGGAGCCATTTTATATCAAGGTT
 GAGATCCATGCTGTAAACAAGACCATGATACCGTCGTCTGTGATCCCAGCCCCAACTATCA
 TCTATTCTCCTGGGGAAG

Intron f/g

GTGAGAGAACCAGTAATAGCTACTGTCTACAAAGAATGTGTTCAATTTAAAGACCTGACTGT
 AGGCCGATGGCTGCTGTCATCTCCTCCGCCTCCTCCTCCTGTTCCCTCCTCCGAAGGGGTCA
 GCTTCAGGTTCTCTTGCCAATATGCCAAGCAGACCTCCTGAGCAGGCAGTATATATACGTA
 AGGGAAGCAAGTATGGACCATCGCGCGGCATGTAGAGATACAATGATCAGCTGTCTGCTGT
 TCCACTCCTGTCAGACAATGAGATAAACATGAATACAGTATTACTCAGCAGCGTTCCAATT
 TTCAACCCCTCGTATTTATTAAAAAAGGAATTTTAAATATATTTTCTCCTTGTTGAAATA
 TTTTAGTAAGTGTAAATCGATATAGAGTGGAGTAGTGACGCTTTATTTCCGGTTCATTCTCG
 AAACAAAAATATAATAGTCCACTGAACTCTCTTAAATTGTTTTTACAACCTTCAACTGCCA
 CAGACGTAATCCCTCACGTTATTTTGAGCTGACAACGTGTTGAATTGAGTGTGTTCCGAAT
 TCTAAATAAGCATGTATATATTTACGTCTCATGCAAGTAATATATGTTTAACTGATGACGT
 CACTTGGTGACCACTGATTTAGTTCCCTTTGTCATAATTGCAGTTTCTGTTGTCACGGGGAC
 GGTGGGGAAGCCAGGTTCCCTCCTGTCACGCTGAATATCCCGTTTCAATCCCCCACATGGGT
 ACAAAGTGTGATGCCTATTTCTGGTGTCCCCACCGTGATATTGCTGGAATAAGTGGCTTA
 ATACCATATACACTCACTCTATTGTCACTACTGCCACCGGCTCACACCTCTGATGCTTC
 TGTTCATCCAG

Domäne g(1)

GTCGCGCTGCTGACAGTGCGCACTCTGCCAACATTGCTGGCTCTGGGGTGAGGAAGGACGT
 CACGACCCTCACTGTGTCTGAGACCGAGAACCTAAGACAGGCTCTTCAAGGTGTCATCGAT
 GATACTGGTCCCAATGGTTACCAAGCAATAGCATCCTTCCACGGAAGTCTCCAATGTGCG
 AGATGAACGGCCGCAAGGTTGCCTGTTGTGCTCAG

Intron g(1)

GTAATTAATGGATGTGAAGTCAATGTCCGAGGGTATAATAAGGATTTAAATACTTCAGTCG
 TGTAATACTGTATGACATGTGTATTGGATGGTGTAGGTATTACAGGTTATAAGGCCAGTGT
 GTGTTGGGACGGTTACTTTCCCTGCACTAGTAATAAGCATTGTATTTAGCTAGCTTTTATCA
 TATACTTTAGTTTCAGGTTTGtGGCAATTGAAATCGAAATTTTCTTTCATTTCAAGGTTA
 TCGCACTCGTGTGTNAGAATAGTTACTATGCTGCATTGAGAATAACACTATAGTAATAAAG
 CATATCATACAGTAAGAATAACACTATAGTAATAAAGTATATCATNCAGTAAGAATGTCAT
 TGTATGATAAATAGGTTATCACACTCGTGTGTTTTAGAATGGTTACTATCCCAGGAATAAC
 CACTATGTATTACATGTATATTGGGCAGTGTAAGTAGTAGCATTGTATATTAAATCAGTAT
 ATCGTGCTTCAAAACACCAGGATATATGGGGTATACAGTGGGCAGTGTAAGTAGCAACATT
 GTATATTAAATCAGTATATCGTACTTCAAAACACCAGGATTATGGGGTATACAGTGGGCAG
 TGTAAGTAGTAGCATTGTATATTAAATCAGTATATCGTACTTCAAAACACCAGGATATAAT
 TCAGTATATCGTGCTTCAAAACACCAGGATATAATTTCAGTATATCGTGCTTCAAAACACCA
 GGATATATGGGATATACAGTGCGGGTTTGCATACAACCTCCACCCTTTACAG

Domäne g(2)

GTATGGcCTCCTTCCcACACTGGCACAGACTGTATGTGAAGCAGATGGAAGATGCCCTGGC
 TGACCACGGGTCACATATCGGCATCCCTTACTGGGACTGGACAACCTGCCTTCACAGAGTTA
 CCCGCCCTTGTACAGACTCCGAGAACAAATCCCTTCCATGAG

Intron g(2)

GTCAGTTTAGTCTCCTGTCTGAGCTAACGATACCAATTTCTATTTTCGAGAACCACGATG
 ACGAGAAAACAAGCAATATAGATATAGATGCAGTATAGATCAAGTTAATGAATTCATTGCT
 ATATGTTTGCTTGTAATAAACTTTAAGAAAACGAGAGCATGCACACAAATGAAACAAACAA
 TTATGTGTTTGATAGGAATATGATATATGTATTTGGGGGCTGACGTGAGCAGGGTTGAAGG
 GACAGTTTACATTGTCAGTAACACTGGGAGTATTCTTTGATCCACAATATATAGTTTCATT
 GTGTTTCAGCAGTTACAACATAACATTATATCATACATTACGTGtAACATGCTTCTTTTGTC
 CTCTTTTGCCAG

Domäne g(3)

GGTCGCATTGATCATCTCGGTGTAACCACGTCACGTTCCCCCAGAGACATGCTGTTTAACG
 ACCCAGAGCAAGGATCAGAGTCGTTCTTCTATAGACAAGTCCTCCTGGCTTTGGAGCAGAC
 TGACTACTGCCAGTTCGAAGTCCAGTTTGAGCTGACCCACAACGCCATTCACTCCTGGACA
 GGTGGACGTAGCCCTTACGGAATGTCGACCCTCGAGTTCACAGCCTACGATCCTCTCTTCT
 GGCTTCACCACTCCAACACCGACAGAATCTGGGCTGTCTGGCAAGCACTGCAGAAATACCG
 AGGACTCCCATAACAACGAAGCACACTGTGAAATCCAGGTTCTGAAACAGCCCTTGAGGCCA
 TTCAACGATGACATCAACCACAATCCAATCACCAAGACTAATGCCAGGCCTATCGATTTCAT
 TTGATTATGAGAGGtTTAACTATCAGTATGACACCCTTAGCTTCCATGGTAAGAGCATCCC
 TGAAGTGAATGACCTGCTCGAGGAAAGAAAAAGAGAAGAGAGAACATTTGCTGcCTTCCTT
 CTTTCGTGGAATCGGTTGCAGTGCTGATGTCGTCTTTGACATCTGCCGgCCCAATGGTGACT
 GTGTCTTTGCAGGAACCTTTGCTGTGCTGGGAGGGGAGCTaGAAATGCCTTGGTCCTTCGA
 CAGACTGTTCCGCTATGACATCACCAGAGTCATGAATCAGCTCCATCTCCAGTATGATTCA
 GATTTTCAGTTTCAGGGTGAAGCTTGTTGCCACCAATGGCACTGAGCTTTCATCAGACCTtC
 TCAAGTCACCAACAATTGAACATGAACCTTG

Intron g/h

GTATGTTATCTTATCATCAAATGTGTGATCAGATACTGGAGACGTTTTTCATATTAACCTGG
 TCAGCATTAGTTGATGATTTTTGGTGCGATGTTGACGACAAGGAGTCAAGCATTAAACACATT
 CAACACATCTTTAATCTGATATGAGAAGGGAATAAATTGATCCAGTATTGATGATTGAAGT
 TAGATTAACAGTGAAAGATATACCAGTTTTGATAATCGTATAAAACAGTAGCAGAATTGTA
 TCGTGAAAACATAAATGTGGGAAGGCCAAGCCAAAGCAGATTTTAGATTACGATCGTGTGCT
 AGAATAATTACAATAACCCAGACGTCGGAATGTGGTTGTCTATGGCAATGGTTACGATT
 AATTGCTAACATGCACGATTTACCTATTTTCAG

Domäne h

AGCCACAGAGGACCAGTTGAAGAAACAGAAGTCACTCGCCAACATACTGACGGCAATGCA
 CACTTTTCATCGTAAGGAAGTTGATTGCTGTCCCTGGATGAAGCAAACAACCTTGAAGAATG
 CCCTTTACAAGCTACAGAACGACCACAGTCTAACGGGATACGAAGCAATCTCTGGTTACCA
 TGGATACCCCAATCTGTGTCCGGAAGAAGGCGATGACAAAATACCCCTGCTGCGTCCCCGG
 ATGGGCATCTTTCCTTACTGGCACAGACTCTTGACCATTCAACTGGAAAGAGCTCTTGAGC
 ACAATGGTGCACCTGCTTGGTGTTCCTTACTGGGACTGGAACAAGGACCTGTCGTCACTGCC
 GCGTTCCTTCTCCGACTCCAGCAACAACAATCCCTACTTCAAGTACCACATCGCCGGTGTT
 GGTACAGACACCGTCAGAGAGCCAACCTAGTCTTATATATAACCAGCCCCAAATCCATGGTT
 ATGATTATCTCTATTACCTAGCATTGACCACGCTTGAAGAAAACAATTACTGGGACTTTGA
 GTTTCAGTATGAGATCCTCCACAACGCCGTCCTCCTGGCTTGGAGGATCCCAGAAGTAT
 TCCATGTCTACCCCTGGAGTATTTCGGCCTTTGACCCTGTCTTTATGATCCTTCACTCGGGTC
 TAGACAGACTTTGGATCATCTGGCAAGAACTTCAGAAGATCAGGAGAAAGCCCTACAACCT
 CGCTAAATGTGCTTATCATATGATGGAAGAGCCACTGGCGCCCTTCAGCTATCCATCTATC
 AACCAGGACGAGTTCACCCGTGCCAACTCCAAGCCTTCTACAGTTTTTGACAGCCATAAGT

17/ 26

TCGGCTACCATTACGATAACCTGAATGTTAGAGGTCACAGCATCCAAGAACTCAACACAAT
CATCAATGACTTGAGAAACACAGACAGAATCTACGCAGGATTTGTTTTGTCAGGCATCGGT
ACGTCTGCTAGTGTCAAGATCTATCTCCGAACAGATGACAATGACGAAGAAGTTGGAACCTT
TCACTGTCCTGGGAGGAGAGAGGGAAATGCCATGGGCCTACGAGCGAGTTTTCAAGTATGA
CATCACAGAGGTTGCAGATAGACTTAAAATTAAGTTATGGGGACACCCTTTAACTTCCGGA
ACTGGAGATCACATCCTTACGAATGGAATCGGTGGTAAACAAGAGCCTACCCAAATCCTTT
CATCATCTACAGACCTGCCAATCATGACTACGATGTTCTTGTATCCCAGTANGGAAGAAA
CCTTCACATCCCTCCCAAAGTTGTCGTCAAGAAAGGCACCCGCATCGAGTTCACCCAGTC
GATGATTCAGTTACGAGACCAGTTGTTGATCTTGGAAAGCTACACTGCACTCTTCAACTGTG
TGGTACCACCGTTCACATACCACGGATTTCGAACCTGAACCACGTCTATTCTGTCAAGCCTGG
TGACTACTATGTTACTGGACCCACGAGAGACCTTTGCCAGAATGCAGATGTCAGGATTCAT
ATCCATGTTGAGGATGAGTAA

3' UTR

CGCAACAGGT

Intron UTR

GAGATAAGAAACCCTTCTAACAGTAATACGACACCACATTACAGCTTAAACATGATTGCCA
TCGATGTTTTTCATGTGTAGTATACGCTTTTCAGTTCTACATAATTTTGTTTTTCAAATCAA
GTTTAGCAAATGAATCTATCACTGGAAAATAGGGTAGGGTAGCCAAGTGGTTAAAGCGGTC
ACTGATCACGCCAAAGACGAGTGTCTAACCTGCATGGGTACAAAAGTGAAGACCATTGCT
GGTGTCTACCGCCGTAATATTGTTTTTAGTATTGCTAAAACTTATACTCACCCATGCGCTG
TAAAAGTGAATAATAATCATATTTCAACAAAAGCACAAAACCATTTTCATTTTCATGAAAG
CCTCTTGTTACCTGAAAGACGCAAGAGAACAATAGTTCCTAACATTATTTTCAGACATTG
GAAATGTCCTGCACGTGTAAACCATATATCCTTTGAAATTTTACGACTGCATCGTATACA
ATTTATGATATAAATTTAAACTTTAT

3' UTR

TTCTTGGTCTCCACATATTCACATATCAGCACCAAATGGTTTCGAAGGACATTGGCGTTCT
TCTCTGGCAATGCATTTCAATACAACATTGAAAATGACTTCAGCATATCAGTGTGCTTCGA
ACGTGTTCCGGAAGTACTCAAATGTGCTATGACTGAATTATTGTACATACATAACTTATTG
ATGTTCAATAAATAAATGTTGAAACGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figur 7

Abgeleitete Primärstruktur des HtH2

Domäne b

HRLFVTQVEDALIRRGSPIGVPYWDWTQPMALPGLADNATYRDPISGDSRHNPFHDVEVA
 FENGRTERHPDSRLFEQPLFGKHTRLFDSIVYAFEQEDFCDFEVQFEMTHNNIHAWIGGGE
 KYSSSLHYTAFDPIFYLRHSNTDRLWAIWQALQIRNRNPYKAHCAWSEERQPLKPPAFSS
 PLNNEKTYENSVPTNVYDYEGLVGYTYDDLNFGGMDLGQLEEYIQRQRQDRTFAGFFLS
 HIGTSANVEIIIDHGTLHTSVGTFAVLGGGEKEMKWGFDRLYKYEITDELRLQLNLRADDVFS
 ISVKVTDVDGSELSSSELIPSAAIIFERSH

Domäne c

IDHQDPHHTIIRKNVDNLTPEEINSLRRAMADLQSDKTAGGFQQIAAFHGEPKWCPSPDA
 EKKFSCCVHGMVFPWHRLTVQGENALRKHGCLGALPYWDWTRPLSHLPDLVLVSSRTT
 PMPYSTVEARNPWYSGHIDTVGVDTTTSVRQELYEAPGFGHYTGVAQVLLALEQDDFCDF
 EVQFEIAHNFIHALVGGSEPYGMASLRYTTYDPIFYLRHSNTDRLWAIWQALQKYRGKPYN
 SANCAIASMRKPLQPFGLTDEINPDDETRQHVPFVSFDYKNNFNYYEYDTLDFNGLSISQL
 DRELSRRKSHDRVFAGFLLHGIQQSALVKFFVCKSDDDCDHYAGEFYILGDEAEMPWGYNR
 LYKYEITEQLNALDLHIGDRFFIRYEAFLHGTSLGSGNIFPKPSVIHDEGA

Domäne d

GHHQADEYDEVVTAASHIRKNLKDLSKGEVESLRS AFLQLQNDGVYENIAKFHGKPGLCDD
 NGRKVACCVHGMPTFPQWHRLYVLQVENALLERGS AVSVPYWDWTETFTELPSLIAEATYF
 NSRQOTFDPNPFFRGKISFENAVTTRDPQPELYVNRYYYQNVMLVFEQDNYCDFEIQFEMV
 HNVLHAWLGGRATYSISSLDYSAFDPVFFLHHANTDRLWAIWQELQRYRKKPYNEADCAN
 LMRKPLHPFDNSDLNHDPTFKYSKPTDGFYQNNFGYKYNLEFNHFSIPRLEEIIIRIRQ
 RQDRVFAGFLLHNICTSATVEIFVCVPTTSGEQNCENKAGTEAVLGGGETEMAFHEDRLYRF
 DISETLRDLGLQLDSDHFDLSIKIQGVNGSYLDPHILPEPSLIFVPGSS

Domäne e

SFLRPDGHSDDILVRKEVNSLTRETASLIHALKSMQEDHSPDGFQAIASFHALPPLCPSP
 SAAHRYACCVHGMATFPQWHRLYTVQFDALRRHGATVGVYWDWLRPQSHLPCLVMTET
 HDIWSNRDFPNPFYQANIEFEGENITTE REVIADKLFVKGGHVFDKLVLTQSHPSAEQENY
 CDFEIQFEILHNGVHTWVGGSRTYSIGHLHYAFYDPLFYLRHFFQTDRIWAIWQELQEQRGL
 SGDEAHCALEQMREPLKPPSFGAPYNWNQLTQDFSRPEDTFDYRKFGYEYDNLEFLGMSVA
 ELDQYIIHQENDRVFAGFLLSGFGGSASVNFQVCRADSTCQDAGYFTVLGGSAEMAWAFD
 RLYKYDITETLEKMHLRYDDDFITISVSLTANNGTVLSSSLIPTPSVIFQRGH

Domäne f

RDINTRSMSPNRVRRELS DLSARDLSSLKSALRDLQEDDGPNGYQALAAFHGLPAGCHDSR
 GNEIACCIHGMPTFPQWHRLYTLQLEMALRRHGSSVAIPYWDWTKPISELPSLFTSPEYYD
 PWHDAVVNNPFSKGFVKFANTYTVRDPQEMFLQCEHGESILYEQTLLALEQTDYCDFEVQ
 FEVLHNVIIHYLVGGRQTYALSSSLHYASYDPFFFIHHSFVDKMWVWQALQKRRKLPYKRAD
 CAVNLMTKPMRPFDSMDNQNPF TKMHAVPNTLYDYETLYSYDNLEIGGRNLDQLQAEIDR

SRSHDRVFAGFLLRGIGTSADVRFWICRNENDCHRGGIIFILGGAKEMPWSFDRNFKFDIT
HVLENAGISPEDVFDAAEPFYIKVEIHAVNKTMI PSSVIPAPTIIYSPGE

Domäne g

GRAADSAHSANIAGSGVRKDVTTTLTVSETENLRQALQGVIDDTGPNGYQAIASFHGSPPMC
EMNGRKVACCAHGMASFPWHRLYVKQMEDALADHGSHIGIPYWDWTTAFTELPALVTDSE
NNPFHEGRIDHLGVTTSRSPRDMLENDPEQGSSEFFYRQVLLALEQTDYQCQFEVQFELTHN
AIHSWTGGRSPYGMSTLEFTAYDPLFWLHHSNTDRIWAVWQALQKYRGLPYNEAHCEIQVL
KQPLRPFNDNDINHNPIKTNARPIDSFYERFNYQYDTLSFHGKSIPELNDLLEERKREER
TFAAFLLRGIGCSADVVDICRPNNGDCVFAGTFAVLGGELEMPWSFDRLFRYDITRVMNQL
HLQYDSDFSFRVKLVATNGTELSSDLLKSPTIEHEL

Domäne h

GAHRGPVEETEVTROHTDGNNAHFHRKEVDSLDEANNLKNALYKLQNDHSLTGYEAI SGY
HGYPNLCPEEGDDKIPLLRPRMGIFPYWHRLTIQLERALEHNGALLGVYPYWDWNKDLSSL
PAFFSDSSNNNPYFKYHIAGVGHDVTREPTSLIYNQPQIHGYDYLYYLAITTEENNYWDF
EVQYEILHNAVHSWLGGSSQKYSMSTLEYSAFDPVFMILHSGLDRLWIIWQELQKIRRPYN
FAKCAYHMMEEPLAPFSYPSINQDEFTRANSPSTVFD SHKFGYHYDNLNVRGHSIQELNT
IINDLRNTDRIYAGFVLSGIGTSASVKIYLRTDDNDEEVGTFTVLGGEREMPWAYERVFKY
DITEVADRLKIKLWGHPLTSGTGDHILTNGIGGKQEQPTQILSSSTDLPIMTTMFLLSQXGR
NLHIPPKVVVKGTRIEFHPVDDSVTRPVVDLGSYTALENCVPPFTYHGFELNHVYSVKP
GDYYVTGPTRDLCQNADVRIHIHVEDE

Figur 8

cDNA-Sequenz in Verbindung mit Intronstruktur des KLH1

Domäne b

GGCCTACCGTACTGGGACTGGACTGAACCCATGACACACATTCCGGGTCTGGCAGGAAACA
 AAACCTTATGTGGATTCTCATGGTGCATCCCACACAAATCCTTTTCATAGTTCAGTGATTGC
 ATTTGAAGAAAATGCTCCCCACACCAAAAGACAAATAGATCAAAGACTCTTTAAACCCGCT
 ACCTTTGGACACCACACAGACCTGTTCAACCAGATTTTGTATGCCTTTGAACAAGAAGATT
 ACTGTGACTTTGAAGTCCAATTTGAGATTACCCATAACACGATTACAGCTTGGACAGGAGG
 AAGCGAACATTTCTCAATGTCGTCCCTACATTACACAGCTTTCGATCCTTTGTTTTACTTT
 CACCATTCTAACGTTGATCGTCTTTGGGCCGTTTGGCAAGCCTTACAGATGAGACGGCATA
 AACCTTACAGGGCCCACTGCGCCATATCTCTGGAACATATGCATCTGAAACCATTTCGCCTT
 TTCATCTCCCCTTAACAATAACGAAAAGACTCATGCCAATGCCATGCCAAACAAGATCTAC
 GACTATGAAAATGTCCTCCATTACACATACGAAGATTTAACATTTGGAGGCATCTCTCTGG
 AAAACATAGAAAAGATGATCCACGAAAACCAGCAAGAAGACAGAATATATGCCGGTTTTCT
 CCTGGCTGGCATACTGTTTTCAGCAAATGTTGATATCTTCATTAAAACCTACCGATTCCGTG
 CAACATAAGGCTGGAACATTTGCAGTGCTCGGTGGAAGCAAGGAAATGAAGTGGGGATTG
 ATCGCGTTTTTCAAGTTTGACATCACGCACGTTTTTGAAGATCTCGATCTCACTGCTGATGG
 CGATTTCGAAGTTACTGTTGACATCACTGAAGTCGATGGAACCTAACTTGCATCCAGTCTT
 ATTCCACATGCTTCTGTCATTTCGTGAGCATGCACGTGGTAAGCTGAATAGAG

Intron b/c

GTTTTGTAATAATTATGTAGAATTCTTTACCTCAGAATAAGATGAGGTCACATGGGTTTTG
 CAAACTATTACGTTTGAATTAATATTAATAATACCGGACCCTCCACTGGTACATATTTAT
 CTTTATAACGATAATAGCGATGATGATGATGATGATGATGATGATGATGATGATgATAATg
 ATgATGCCGGTATTGCACGTAATCCAGCCGAcTTAGATGACACCCTAAGGGTGCAGAAAGT
 ATAaCAATTAGATTGCGTTtGCATCTGTGTATGCGTGTGCTTTaCCAAAAGTCAAAATAA
 AAGTGCAAACCTTAGTTTATTCAATTTGATAGAGCCTTTTACGATAAGAACAATGTAATAA
 ATTAGAACATAACTGAAACCTCCGAAAGAAGGCCTGTTTGTCAAGAGAGGTATCGACATGA
 TTGACTTATAAACCTGTGCTTCTATATTTTGGAACTGTCCACTTTCTTGTGTGTGTACTG
 TAATCACATCGCACTATGGCTGCAAGACGTGTACGAGTACACTATATACTTACCTAATGAC
 CAACCACAAGGCTGGCTTTGTTAATATTGTTATTTTACAGAAATAAACACAGAATTCCAGC
 ATTTGGCTGGTGTATTTAGCAAAACACCGATATGACACTCATGTTTTATTACATTTTTTTTC
 AG

Domäne c

TTAAATTGACAAAGTGCCAAGGAGTCGTCTTATTTCGAAAAAATGTAGACCGTTTGAGCCC
 CGAGGAGATGAATGAACTTCGTAAAGCCCTAGCCTTACTGAAAGAGGACAAAAGTGCCGGT
 GGATTTTCAGCAGCTTGGTGCATTCCATGGGGAGCCAAAATGGTGTCTTAGTCCCGAAGCAT
 CTAAAAAATTTGCCTGCTGTGTTTACGGCATGTCTGTGTTCCCTCACTGGCATCGACTGTT
 GACGGTTCAGAGTGAAAATGCTTTGAGACGACATGGCTACGATGGAGCTTTGCCGTACTGG
 GATTGGACCTCTCCTCTTAATCACCTTCCCGAACTGGCAGATCATGAGAAGTACGTCGACC
 CTGAAGATGGGGTAGAGAAGCATAACCCTTGGTTCGATGGTCATATAGATACAGTCGACAA
 AACAACAACAAGAAGTGTTCAGAATAAACTCTTCGAACAGCCTGAGTTTGGTCATTATACA
 AGCATTGCCAAACAAGTACTGCTAGCGTTGGAACAGGACAATTTCTGTGACTTTGAAATCC
 AATATGAGATTGCCATAACTACATCCATGCACTTGTAGGAGGCGCTCAGCCTTATGGTAT
 GGCATCGCTTCGCTACACTGCTTTTGATCCACTATTCTACTTGCATCACTCTAATACAGAT
 CGTATATGGGCAATATGGCAGGCTTTACAGAAGTACAGAGGAAAACCGTACAACGTTTGCTA
 ACTGTGCTGTTACATCGATGAGAGAACCCTTTGCAACCATTTGGCCTCTCTGCCAATATCAA
 CACAGACCATGTAACCAAGGAGCATTCAGTGCCATTCAACGTTTTTGTATTACAAGACCAAT

TTCAATTATGAATATGACACTTTGGAATTTAACGGTCTCTCAATCTCTCAGTTGAATAAAA
 AGCTCGAAGCGATAAAGAGCCAAGACAGGTTCTTTGCAGGCTTCCTGTTATCTGGTTTCAA
 GAAATCATCTCTTGTAAATTCAATATTTGCACCGATAGCAGCAACTGTCACCCCGCTGGA
 GAGTTTTACCTTCTGGGTGATGAAAACGAGATGCCATGGGCATACGATAGAGTCTTCAAAT
 ATGACATAACCGAAAACTCCACGATCTAAAGCTGCATGCAGAAGACCACTTCTACATTGA
 CTATGAAGTATTTGACCTTAAACCAGCAAGCCTGGGAAAAGATTTGTTCAAGCAGCCTTCA
 GTCATTTCATGAACCAAGAATAG

Intron c/d

GTAATTGTTATATGTTTCGAATATTGCCGATACCTTCAATATATATACTTTATCAAAGTAA
 TTGATTAATCTGAAGTAATTTTCCTTTCCAGTAGAGATTCAGTTGATACAACAAGAATTCG
 CCCTGTTGTATGTCACCTTATTTTCATCAAACGATTTCGAAGTGAGCTGTCCATGCCACAAT
 GGGGTCTCTGTAACCTTTCTCGTATGGGGTATAGATTATATAGACGTGGCAGACCTTACGTA
 TAACTAATATTTGTGTAATGTCGTTTCAG

Domäne d

GTCACCATGAAGGCGAAGTATATCAAGCTGAAGTAACTTCTGCCAACCGTATTCGAAAAAA
 CATTGAAAATCTGAGCCTTGGTGAACCTCGAAAGTCTGAGAGCTGCCTTTCCTGGAAATTGAA
 AACGATGGAACCTTACGAATCAATAGCTAAATTCCATGGTAGCCCTGGTTTGTGCCAGTTAA
 ATGGTAACCCCATCTCTTGTGTGTCCATGGCATGCCAACTTTCCCTCACTGGCACAGACT
 GTACGTGGTTGTCTGTTGAGAATGCCCTCCTGAAAAAAGGATCATCTGTAGCTGTTCCCTAT
 TGGGACTGGACAAAACGAATCGAACATTTACCTCACCTGATTTTCAGACGCCACTTACTACA
 ATTCCAGGCAACATCACTATGAGACAAACCCATTCCATCATGGCAAAATCACACACGAGAA
 TGAAATCACTACTAGGGATCCCAAGGACAGCCTCTTCCATTTCAGACTACTTTTACGAGCAG
 GTCCTTTACGCCTTGGAGCAGGATAACTTCTGTGATTTTCGAGATTCAGTTGGAGATATTAC
 ACAATGCATTGCATTCTTTACTTGGTGGCAAAGGTAAATATTCCATGTCAAACCTTGATTA
 CGCTGCTTTTGATCCTGTGTTCTTCTTCATCACGCAACGACTGACAGAATCTGGGCAATC
 TGGCAAGACCTTCAGAGGTTCCGAAAACGGCCATACCGAGAAGCGAATTGCGCTATCCAAT
 TGATGCACACGCCACTCCAGCCGTTTGATAAGAGCGACAACAATGACGAGGCAACGAAAAC
 GCATGCCACTCCACATGATGGTTTTGAATATCAAACAGCTTTGGTTATGCTTACGATAAT
 CTGGAACCTGAATCACTACTCGATTCTCAGCTTGATCACATGCTGCAAGAAAGAAAAAGGC
 ATGACAGAGTATTCGCTGGCTTCCTCCTTCACAATATTGGAACatCTGCCGATGGCCATGT
 ATTTGTATGTCTCCCAACTGGGGAACACACGAAGGACTGCAGTCATGAGGCTGGTATGTTT
 TCCATCTTAGGCGGTCAAACGGAGatGTTCCTTTGTATTTGACAGACTTTACAAACTTGACA
 TAACTAAAGCCTTGAAAAAGAAGGTTGTGCACCTGCAAGGGGATTTTCGATCTGGAAATTGA
 GATTACGGCTGTGAATGGATCTCATCTAGACAGTCATGTCATCCACTCTCCCACTATACTG
 TTTGAGGCCGGAACAG

Intron d/e

GTAACATTTTTGTCACTGTAACCAACAACCTGCAGTCTATTTTGCAATTACGATAATAACAA
 TTTTTGAAATATATCTTTATTAAAGCAAAGGTTTCTAGAGACAAACAGCCGGCTCTAATTA
 TTTTTTCGAACTTACGCTTGAGTAAAGATCTGCAAATGGCAACCCTACCTATACTATTAA
 AATATAATGTTACATTCGTATCTGAATGTTTAATAAATCACTTCATATTCTGTTGCAG

Domäne e

ATTCTGCCCCACACAGATGATGGACACACTGAACCAGTGATGATTCGCAAAGATATCACACA
 ATTGGACAAGCGTCAACAACCTGTCACTGGTGAAAGCCCTCGAGTCCATGAAAGCCGACCAT
 TCATCTGATGGGTTCCAGGCAATCGCTTCCTTCCATGCTCTTCCTCCTCTTTGTCCATCAC
 CAGCTGCTTCAAAGAGGTTTGCGTGCTGCGTCCATGGCATGCCAACCTTCCCGCAATG

Figur 9

Abgeleitete Primärstruktur des KLH1

Domäne b

GLPYWDWTEPMTHIPGLAGNKTYVDSHGASHTNPFHSSVIAFEENAPHTKRQIDQRLFKPA
 TFGHHTDLFNQILYAFEQEDYCDFEVQFEITHNTIHAWTGGSEHFSMSSLHYTAFDPLFYF
 HHSNVDRLWAVWQALQMRHHPYRAHCAISLEHMLKPPAFSSPLNNNEKTHANAMPNKIY
 DYENVLHYTYEDLTFFGGISLENIEKMIHENQQEDRIYAGFLLAGIRTSANVDIFIKTDSV
 QHKAGTFAVLGGSKEMKWGFDRVFKFDITHVLKDLDLTADGDFEVTVDITEVDGTKLASSL
 IPHASVIREHARGKLN

Domäne c

VKFDKVPRSRRLIRKNVDRLSPEEMNELRKALALLKEDKSAGGFQQLGAFHGEPKWCPSP
 SKKFACCVHGMSVFPWHRLTLVQSENALRRHGYDGPYWDWTSPLNHLPELADHEKYVD
 PEDGVEKHNPWFDGHIDTVDKTTTRSVQNKLFQEPFGHYTSIAKQVLLALEQDNFCDFEI
 QYEIAHNYIHALVGGAQPYGMASLRYTAFDPLFYLHHSNTDRIWAIWQALQKYRGKPYNVA
 NCAVTSMREPLQPFGLSANINTDHVTKEHSVPFNVDYKTNFNIEYDTLEFNGLSISQLNK
 KLEAIKSQDRFFAGFLLSGFKKSSLVKFNICTDSSNCHPAGEFYLLGDENEMPWAYDRVFK
 YDITEKLHDLKLHAEDHFYIDYEVFDLKPASLGKDLFKQPSVIEPRI

Domäne d

GHHEGEVYQAEVTSANRIRKNIEENLSLGELESLRAAFLEIENDGTYESIAKFHGSPGLCQL
 NGNPISCCVHGMPFPHWHRLYVVVENALLKKGSSVAVPYWDWTKRIEHLPHLISDATY
 NSRQHYYETNPFHGGKIITHENEITTRDPKDSLFSYDYFYEQVLYALEQDNFCDFEIQLEIL
 HNALHSLGKGKYSMSNLDYAAFDPVFFLHATTDRIWAIWQDLQRFKRPYREANCAIQ
~~LMHTPLQPFDKSDNDEATKTHATPHDGFYQNSFGYAYDNLELNHYSIPQLDHMLQERKR~~
 HDRVFAGFLLHNIGTSADGHVFCVCLPTGEHTKDCSHEAGMFSILGGQTEMSFVFDRLYKLD
 ITKALKKNGVHLQGDFFLEIEITAVNGSHLDHSHVHSPTILFEAG

Domäne e

DSAHTDDGHTPEVMIRKIDITQLDKRQQLSLVKALESMAKADHSSDGFQAIASFHALPPLCPS
 PAASKRFACCVHGMPFPHWHRLYTVQFQDSLKRGAVVGLPYWDWTLPR

Figur 10

cDNA-Sequenz in Verbindung mit Intronstruktur des KLH2

Domäne b

GGCCTGCCCTACTGGGATTGGACCATGCCAATGAGTCATTTGCCAGAACTGGCTACAAGTG
 AGACCTACCTCGATCCAGTTACTGGGGAACTAAAAACAACCCTTTCCATCACGCCCCAAGT
 GGCGTTTGAAAATGGTGTAAACAAGCAGGAATCCTGATGCCAAACTTTTTATGAAACCAACT
 TACGGAGACCACACTTACCTCTTCGACAGCATGATCTACGCATTTGAGCAGGAAGACTTCT
 GCGACTTTGAAGTCCAATATGAGCTCACGCATAATGCAATACATGCATGGGTGGAGGCAG
 TGAAAAGTATTCAATGTCTTCTCTTCACTacacTGCTTTTGATCCTATATTTTACCTCCAT
 CACTCAAATGTTGATCGTCTCTGGGCCATTTGGCAAGCTCTTCAAATCAGGAGAGGCAAGT
 CTTACAAGGCCCACTGCGCCTCGTCTCAAGAAAGAGAACCATTAAAGCCTTTTGCATTTCAG
 TTCCCCACTGAACAACAACGAGAAAACGTACCACAACCTCTGTCCCCACTAACGTTTATGAC
 TATGTGGGAGTTTTGCACTATCGATATGATGACCTTCAGTTTGGCGGTATGACCATGTCAG
 AACTTGAGGAATATATTACAAAGCAGACACAACATGATAGAACCTTTGCAGGATTCTTCCT
 TTCATATATTGGAACATCAGCAAGCGTAGATATCTTCATCAATCGAGAAGGTCATGATAAA
 TACAAAGTGGGAAGTTTTGTAGTACTTGGTGGATCCAAAGAAATGAAATGGGGCTTTGATA
 GAATGTACAAGTATGAGATCACTGAGGCTCTGAAGACGCTGAATGTTGCAGTGGATGATGG
 GTTCAGCATTACTGTTGAGATCACCGATGTTGATGGATCTCCCCCATCTGCAGATCTCATT
 CCACCTCCTGCTATAATCtTTGaACGTGGTCaTG

Intron 2b/c

AGGTATTTAAAAAAGTAATAAAACCcATATTTTCGAATGCGCTTTATGAAATATCGTGTGAC
 TGTTTCTTTAGTTTACATGGAGTGTAACAACATGCTCCATCAGTTGACATATACTGCTCAC
 ACAAGTAAGGGATATTTGATAATGATAACAAATATAATCAAAGCGGTTATACTATCAAGA
 CTTATTCACATAATTACAGGTGAAGGGAGGTGTGATCGTGTTCACTGATCAGGTTGAGGCC
 AGAGAAGTCCCAGTTTGAGTCTTGCAAGAGATGATGTTTAGGCATGGGGTCAATCACCAA
 AATCACATGACTTCAATAACGGGTTGGACCACCTCGAGCGACgATGCAAGCAGTAGAGCGT
 CTACGCATGCTCTGATAAGGCGACCAATCTGTTCTGGGGAATCAGtCGCCACTCCTCTT
 GTAGTGCCACGCTCATTCTGCTACGGTCTGGGTACCTGCTATCGGgTCTTGATCCGTAT
 CCCAAGGATGTCCACACATGTTCAAgGTGAGAGGTGGGGAACATCGCTGGCCACGGTaA
 GGtCTGAATTTGATGCCGTTGAAAGTGAGCTCTGACAACcTGAGCATGGtGAGCTCTGACG
 TTGTCGTCCTGAAAGATGAATcCAGCTcCaTGaCAGCGAGCAAaGGGCAGGACGTGTTGGT
 CAATGCAGTTGTCTCTGCAGTACACACCTGTCACTCGCCACTCACAAGCGTGTAGATCTGT
 ACGACCAGTCATGGAGATCCCAGCCACATCATAACGGACCCCTATCCATACCGATCATGA
 GCCACCATAGCAGCGTCTTGATGACGTTCTCCCTGTGCGCTCGACATCCTcACACGGCCAA
 AAGGAACGTGGACTCGTCACTGAACATGACATTAGCCAACCTGGCACTTGTCCACCGCTGA
 TGTGGCGAGACCATTCCAGTCGAGCTCTTCGGTGTCTGGCTTTCATCGATAACACGACGT
 AAGGTCTGCGGGCGTGCAAGACGGCTCTATGCAGGCGATTTTCGGATTGTCTGGGTGCTAAC
 TCTGATCCCAGGTGCCTGCTGAAGTTGATGCTGGATCTGTGTGGCATTGAGATGGCGATTCT
 CTTAGGACTGTGGAGATGATGAATCGATCTTGACTTATGGTGGTGACATTAGGACGTCGGG
 TTCGTGTCTCTATCCTGCACTCTTCCAGTTGTTTCGGTGACGCTCTGGTACCCGGCTGATTAC
 TGACTGAGAATATCCATCTGCCGTGACATGAGCCCTGTGTTGGCCAGCCTGAAGCATTG
 CAATCGCCAGAGACGCTCTTCAAAAGTCATTCGACGCATGGtTTTCTGTTcACAAATGACA
 GCGTAAACAGtTTTTGGtGCTTTTATGCTTCCCAAGAGCATGAAAAACAGTTCATgGG
 TCGtGCACACCTTACATGACAAGtGtGAAAAGtGACTTGcACCCCTTGtGtGTTCCGATG
 CACACTCTGTTTACGTACTGATGCGATTTGGCGTCTAAACATGTTTTGGCGTCTAAACATG
 TTTTCTGTCATGATTCATATACTATTTTGTTCATATTCCTGGCATCAAACCAAACCTACAGTG
 AAATATATTTCAATATCCCCTACTTTGTGTGAGTAGTATAGATCACTGCAGACAACATATA

GACAAtGCAGtTaCaCCGTCAACAATCCCAGTCATTAATTATGATGaCaCTTCCACACATA
 GGTGTGAGTGAATTGTAATTCAaCTGTACACACTTTTCCCGTGAACATTGAGGATCTATATGA
 CTAATATATAACATTAGTATACGTGCAGTTTTGTATCGCTACGACATTGTTGTAACCTCTT
 TGTTTAATCATTTaACAG

Domäne c

CTGATGCCAAAGaCTTTGgCCATAGCAGAAAAATCAGgAAAGcCGTTGATTcTcTGACAGT
 CgAAGAACAaaCTTCGTTGAGgCGAGcTATGgCAGATcTACAGGACGACAAAACATCAGGG
 GGTTCAGCAGATTGCAGCATTCCACGGAGAACCAAAATGGTGTCCAAGCCCCgAAGCGG
 AGAAAAAATTTGCATGCTGTGTTTCATGGAATGGCTGTTTTCCCTCACTGGCACAGATTGCT
 GACAGTTCAAGGAGAAAAATGCTCTGAGGAAACATGGCTTTACTGGTGGACTGCCCTACTGG
 GACTGGACTCGATCAATGAGCGCCCTTCCACATTTTGTGCTGATCCTACTTACAATGATG
 CTATTTCCAGCCAGGAAGAAGATAACCCATGGCATCATGGTCACATAGACTCTGTTGGGCA
 TGATACTACAAGAGATGTGCGTGATGATCTTTATCAATCTCCTGGTTTTCGGTCACTACACA
 GATATTGCACAACAAGTCCTTCTGGCCTTTGAGCAGGACAGTTTCTGTGATTTTGAGGTAC
 AATTTGAAATTGCCATAATTTTACATGCACTGATTGGTGGTAACGAACCATACAGTAT
 GTCATCTTTGAGGTATACTACATACGATCCAATCTTCTTCTTGCACTACTCCAGTACAGAC
 CGACTTTGGGCCATCTGGCAAGCAATCACTAGTGCGGCCGCCTGCAGGTGCACCATAAGGG
 AGAGCTCCCAACGCgtTGGAtGCAATCT

Domäne g

ATGGCTGTGTTTCCGCACTGGCACAGACTGTTTGTGAAACAGATGGAGGACGCACTTGCTG
 CTCATGGAGCTCATATTGGCATACCATACTGGGATTGGACAAGTGCCTTTAGTCATCTGCC
 CGCCCTAGTGACTGACCACGAGAACAATCCCTTCCACCAC

Intron g(2)

GTATGTGTCAAATCGTTTTAGGAACTGCCTTATCCATTTTACAATTACGAGTACAAAATGA
 AAACGGAACTGTGTGACCTCGAAAAGTGCAATCTTTAAAGGATGCAATGTACACAATAAA
 ATGCTCCGATCAAAGCGATGGCTAGAAATCATTTTCCCTCTAATTCCTTTTACACAGCT
 CGGTTCTTTTTAAGTAGGAACAAGTCTCTGCAAAAACATCACAATAAAGAGAACACAGAA
 AAAACCTCATTCTCGTTTCTGTATTCCGAAAATGAAATTTACAATTTCTTTCATTTATAG

Domäne g

GGCCATATTGGTCATCTGAATGTGGATACATCTCGATCTCCAAGAGACATGCTGTTTAATG
 ATCCTGAACAAGGCTCAGAATCATTCTTCTACAGACAGGTTCTCTTGACTCTAGAACAGAC
 AGACTTCTGCCAATTTGAAGTTCAGTTTGAACCTTACACACAATGCCATCCACTCTTGGACT
 GGAGGACATACTCCATATGGAATGTCATCACTGGAATATACAGCATATGATCCACTCTTTT
 ATCTCCACCATTCCAACACTGATCGTATCTGGGCCATCTGGCAGGCACCTCCAGAAATATAG
 AGGTCTTCCATACAACGCAGCTCACTGCGATATccaagttctgaaacaacctcTTAAACCA
 TTCAGCGAGTCCAGGAATCCAAACCCAGTCAACAGAGCCAATTCTAGGGCCGTTGATTCAT
 TTGATTATGAGAAATTCAATTATCAATATGACACACTTACCTTCCACGGACTTTCTATCCC
 AGAACTTGATGCCATGCTTCAAGAGAGAAAGAAGGAAGAGAGAACATTTGCAGCCTTCCTG
 TTGCACGGATTTGGCGCCAGTGCTGATGTTTCGTTTGATGTCTGCACACCTGATGGTCATT
 GTGCCTTTGCTGGAACCTTCGCGGTACTTGGTGGGGAGCTTGAGATGCCCTGGTCCTTTGA
 AAGATTGTTCCGTTACGATATCACAAGGTTCTCAAGCAGATGAATCTTCACTATGATTCT
 GAGTTCCACTTTGAGTTGAAGATTGTTGGCACAGATGGAACAGAACTGCCATCGGATCGTA
 TCAAGAGCCCTACCATTGAACACCATGGAGGAG

Intron g/h

GTATGTTTTGAGATCCACATAATCTTCTACCCCTGTCTCATTTCTAATGCTCTTCAATACAC
AATTTATATAGCCTTTGAGCTTCAGATGTATTACGGACAGGCATTACAGTATACATGTAAT
ATGGTTTTCTGCTATTTGCAAAAATTGTGTCTTATCTCTGTTTCAGATCATCATGGCGGTGA
CACCTAG

Domäne h

GTCACGATCACAGTGAACGTCACGATGGATTTTTTCAGGAAGGAAGTCGGTTCCTGTCCCT
GGATGAAGCCAATGACCTTAAAAATGCACTGTACAAGCTGCAGAATGATCAGGGTCCCAAT
GGATATGAATCAATAGCCGGTTACCATGGCTATCCATTCCTCTGCCCTGAACATGGTGAAG
ACCAGTACGCATGCTGTGTCCACGGAATGCCTGTATTTCCACATTGGCACAGACTTCATAC
AATCCAGTTTGAGAGAGCTCTCAAAGAACATGGTTCTCATTTGGGTCTGCCATACTGGGAC
TGGAC

Figur 11

Abgeleitete Primärstruktur von KLH2

Domäne b

GLPYWDWTMPMSHLPELATSETYLDPVTGETKNNPFHHAQVAFENGVTSRNPDAKLFMKPT
 YGDHTYLFDSMIYAFEQEDFCDFEVQYELTHNAIHAWVGGSEKYSMSSLHYTAFDPIFYLH
 HSNVDRLWAIWQALQIRRGKSYKAHCASSQEREPLKPFASFSSPLNNNEKTYHNSVPTNVYD
 YVGVLHYRYDDLQFGGMTMSELEEYIHKQTOHDRTFAGFFLSYIGTSASVDIFINREGHDK
 YKVGSFVVLGGSKEMKWGFDRMYKYEITEALKTLNVAVDDGFSITVEITDVDGSPPSADLI
 PPPAIIFERGA

Domäne c

DAKDFGHSRKIRKAVDSLTVEEQTSRRAMADLQDDKTSGGFQQIAAFHGEPKWCPSPEAE
 KKFACCVHGMVAFPHWHRLTVQGENALRKHGFTGGLPYWDWTRSMSALPHFVADPTYNDA
 ISSQEEDNPWHHGHIDSVGHDTTRDVRDDLQSPGFGHYTDIAQQVLLAFEQDSFCDFEVQ
 FEIAHNFIHALIGGNEPYSMSSLRYTTYDPIFFLHHSSTDRLWAIWQALQKYRGKPYNTAN
 CAIASMRKPLQPFGLDSVINPDDETREHSVPFRVFDYKNNFDYYESLAFNGLSIAQLDRE
 LQRRKSHDRVFAGFLLHEIGQSAKHNVSDCDHYAGEFYILGDEAEMPWRYDRVYKYEITQQ
 LHDLDLHVGDNFFLKYEAFDLNNGSLGGSIFSQPSVIFEPAAGMF

Domäne d

GSHQADEYREAVTSASHIRKNIRDLSERGEIESIRSAFLQIQKEGIYENIAKFHGKPGLCHE
 DGHVPACCVHGMPTFPHWHRLVVLQVENALLERGSAAVAVPYWDWTLPR

Domäne g

MAVFPHWHRLFVKQMEDALAAHGAHIGIPYWDWTSAFSHLPALVTDHENNPFFHHGHIGHLN
 VDTSRSPRDMLFNDPEQGSSEFFYRQVLLTLEQTDFCQFEVQFELTHNAIHSWTGGHTPYG
 MSSLEYTAYDPLFYLHHSNTDRIWAIWQALQKYRGLPYNAAHCDIQVLKQPLKPFSESARNP
 NPVTRANSRAVDSFDYEKFNYQYDTLTFHGLSIPELDAMLQERKKEERTFAAFLHGFAS
 ADVSFDVCTPDGHCAFAGTFAVLGGELEMPWSFERLFRYDITKVLKQMNLYDSEFHFELK
 IVGTDGTELPSPDRIKSPTIEHHGG

Domäne h

GHDHSERHGDFRKEVGSLSLDEANDLKNALYKLQNDQGPNGYESIAGYHGYPFLCPEHGE
 DQYACCVHGMVFPFPHWHRLHTIQFERALKEHGSHLGLPYWDW